TAX POLICY CENTER
URBAN INSTITUTE & BROOKINGS INSTITUTION

MEASURING SUCCESS: NEW PERFORMANCE METRICS FOR A NEW INTERNAL REVENUE SERVICE

Janet Holtzblatt

September 19, 2024

ABSTRACT

The Inflation Reduction Act (IRA) of 2022 enacted an $80 billion 10-year investment in the Internal Revenue Service on top of the agency's regular annual appropriations. Although Congress cut the IRA funding by over 25 percent the following year, the agency has maintained its commitment to the transformative *IRS Inflation Reduction Act Strategic Operating Plan* to bring the agency into the 21st century. That plan includes detailed objectives and a summary of what success would look like for each new initiative. But it does not provide specific metrics for evaluating the agency's performance in achieving many of those goals. Well-defined performance metrics will provide insight into the overall return on the now nearly $60 billion investment, inform the IRS's choices between various programs, and identify ways to improve a program's effectiveness. The paper lays out a holistic approach to measuring the IRS's performance with outcome measures that are (1) transparent and consistent with the service, enforcement, and equity goals emphasized in the IRS mission statement; (2) tied to measurable outputs and, in some cases, inputs and efficiency; and (3) reflective of the reality that some factors are beyond the control of the IRS. Importantly, performance metrics should be bundled together to allow lawmakers and administrators to understand the trade-offs between goals and weigh the choices between various IRS activities. Finally, the effectiveness of all IRS actions cannot be reduced to a single quantitative metric or even a bundle of metrics. Performance measures are not a replacement for a thorough evaluation of the IRS's actions.

ABOUT THE TAX POLICY CENTER

The Urban-Brookings Tax Policy Center aims to provide independent analyses of current and longer-term tax issues and to communicate its analyses to the public and to policymakers in a timely and accessible manner. The Center combines top national experts in tax, expenditure, budget policy, and microsimulation modeling to concentrate on areas of tax policy that are critical to future debate.

Copyright © 2018. Tax Policy Center. Permission is granted for reproduction of this file, with attribution to the Urban-Brookings Tax Policy Center.

CONTENTS

ABSTRACT II

CONTENTS III

ACKNOWLEDGMENTS V

INTRODUCTION 1

THE INFLATION REDUCTION PLAN 4

Table 1 5

STRATEGIC OPERATING PLAN 7

2023 Strategic Operating Plan 7

2024 Update to Strategic Operating Plan 8

LEGISLATIVE REQUIREMENTS FOR MEASURING THE IRS'S PERFORMANCE 9

Government Performance and Results Act 9

Paperwork Reduction Act 10

Improper Payments Information Act 11

PERFORMANCE MEASURES AND THE IRS MISSION 12

Purpose 12

Baseline 13

TAXPAYER SERVICES 15

Outcome: Taxpayer Satisfaction. 15

Outcome: Compliance Burdens 16

Output: Telephone Service 19

ENFORCEMENT 22

Outcome: Tax Gap 22

Output: Audit Rates 25

Table 2 25

Efficiency: Returns on Investment 27

EQUITABLE TAX ADMINISTRATION 31

Outcome: Distributing Compliance Burdens by Income 31

Outcome: Distributing Noncompliance by Income 31

Output: Distributing Audit Rates by Race 33

CONCLUSION 35

APPENDICES

Appendix 1 38

CONTENTS

Table A1 38

Appendix 2 **40**

Table A2 40

Appendix 3 **46**

Table A3 46

Appendix 4 **49**

Table A4 49

NOTES **52**

REFERENCES **54**

ACKNOWLEDGMENTS

This report was funded by an anonymous foundation. We are grateful to them and to all our funders, who make it possible for the Urban-Brookings Tax Policy Center to advance its mission.

The author thanks Tracy Gordon, Arnstein Ovrum, and attendees at the 14th Annual IRS/TPC Joint Research Conference on Tax Administration for their helpful comments on an earlier draft and Lillian Hunter for assistance in compiling the report. The views expressed are those of the authors and should not be attributed to the Urban-Brookings Tax Policy Center, the Urban Institute, the Brookings Institution, their trustees, or their funders. Funders do not determine research findings or the insights and recommendations of our experts. Further information on Urban's funding principles is available at https://www.urban.org/about/organizational-principles; further information on Brookings' donor guidelines is available at https://www.brookings.edu/support-brookings/donor-guidelines/.

Copyright © 2024. Tax Policy Center. Permission is granted for reproduction of this file, with attribution to the Urban-Brookings Tax Policy Center.

INTRODUCTION

In April 2023, the Internal Revenue Service took a major step toward modernization by releasing the *IRS Inflation Reduction Act Strategic Operating Plan* detailing how the agency will invest the 10-year $80 billion boost to its budget provided under the 2022 Inflation Reduction Act (IRS 2023b). Although that funding was cut by over 25 percent a month later in the Fiscal Responsibility Act, the IRS has maintained its commitment to the plan with the understanding that funding for some initiatives will run out much sooner than initially anticipated.

The plan is a serious and comprehensive effort to bring the agency into the 21st century, but lingering questions must be addressed to ensure its success. First, what is the long-term plan? The remaining $58.6 billion budget-boost is a 10-year investment, but most of the Strategic Operating Plan, as well as a 2024 update, provides details for just the next few years (IRS 2023b, 2024d). Many features of the plan involve recruitment, research, evaluation, and pilot programs—the first steps toward development and implementation of effective long-term strategies.

Second, how will success be measured? The strategic plan contains objectives and a summary of what success would look like for each new initiative. Although the update lists “outcomes,” it does not provide specific metrics or targets for evaluating the agency’s performance in achieving many of the plan's goals—either for the specific initiatives or for the entire plan once fully implemented.

Given the early stages of the plan's implementation, holding the IRS to tough metrics now would be premature. Establishing targets too soon could further discourage efforts to test different approaches to determine which is the most efficient and fair to taxpayers.

But identifying and designing serious performance metrics should begin before the IRS proceeds too far in implementing the various initiatives. This would allow development of serious metrics reflecting thoughtful and careful analysis and in coordination across IRS divisions and with input from outside experts.

In this paper, I establish several principles for designing metrics, outlined below:

- The goals of the performance measures should be transparent. A goal of a metric may be to support the IRS’s budget request for overall funding, another set of metrics could inform internal decisions as to how to best allocate appropriations across programs, and a third set could help the IRS refine and improve a program.
- The metrics should be consistent with the outcomes emphasized in the IRS mission statement: **taxpayer services, enforcement, and equitable treatment of taxpayers**.
- Within those three broad outcome categories, the IRS should measure *output* metrics, which measure the IRS’s achievement of specific actions—such as the number of phone calls answered or the successful targeting of audits.

- In some instances, it may be helpful to establish *input* metrics—for example, the costs of implementation of certain activities—or *efficiency* metrics, such as the return on investment.
- A methodology should be developed to distinguish between the IRS's role in administering the tax code from factors that are beyond the control of the IRS—such as economic conditions, the tax code, and the agency's funding.

Importantly, the metrics should not be viewed in isolation. Some metrics should be bundled together so that policymakers and administrators can assess the trade-offs between goals and weigh the choices between activities—whether it is a choice between services and enforcement, answering phones and opening the mail, and correspondence and in-person audits.

Finally, the effectiveness of all IRS actions cannot be reduced to a single quantitative metric or even a bundle of metrics. Performance measures are not a replacement for a thorough evaluation of the IRS's actions.

In this paper, I review the objectives of the IRS Strategic Operating Plan as well as prior legislation that has required the agency and other government agencies to set performance measures for at least some of its activities. I identify the shortfalls in the current patchwork of metrics and present a more holistic perspective on measuring the performance of the IRS. I then evaluate several examples of outcome, output, and efficiency metrics, pointing to ways those measures could be refined or expanded to provide more insight into the performance of the agency in achieving its mission of providing taxpayer services, enforcing the tax code, and treating taxpayers equitably. The metrics include the following elements:

- Taxpayer services
    1. Taxpayer satisfaction (outcome)
    2. Compliance burdens (outcome)
    3. Telephone service (output)
- Enforcement
    1. Tax gap (outcome)
    2. Audit rates (output)
    3. Return on investment (efficiency)

- Fairness
    1. Compliance burden by income (outcome)
    2. Underreported income and taxes by income group (outcome)
    3. Audit rates by race (output)

I do not discuss the establishment of targets for performance measures. The IRS is typically required to set targets for future performance, but the methodology for deriving those goals is rarely (if ever) described in IRS documents and studies. Lifting the veil on the methods used to set targets would provide more insight into the choice and design of the performance measures.

.

THE INFLATION REDUCTION ACT

In May 2021, the Treasury Department released a multifaceted plan to reduce tax noncompliance. As detailed in "The American Families Plan Compliance Agenda" (US Treasury Department 2021), the impetus was the 19 percent reduction in the IRS's budget, after adjusting for inflation, from fiscal years 2010 through 2020. Those funding reductions contributed to a 20 percent reduction in the IRS workforce and the depreciation of the agency's technological infrastructure.

A key component of Treasury's agenda was a proposal to increase the IRS budget by $80 billion over the next decade. The agenda provided a broad overview of how the IRS would use those funds. Most of the proposed funding would be dedicated to increasing audits of large corporations, partnerships, and global high-wealth and high-income taxpayers. In addition, the massive budget infusion would finance investments in modern technology and new data analytical tools to help select returns for enforcement actions. Finally, the IRS would take steps to improve taxpayer services and facilitate claims of refundable tax credits.

Beyond the unprecedented magnitude of the proposed budget boost, the agenda was unique for two other reasons: First, the funding would cover an entire decade in contrast to the one-year funding typically provided through annual appropriations. Providing a 10-year stream of funds was intended to facilitate long-term planning and investment in technology and staff. Second, the $80 billion boost was conceived as a supplement on top of the annual appropriation—that is, the IRS would still receive funds through the annual appropriations legislation for its normal operating expenses.

In August 2022, Congress passed the Inflation Reduction Act (IRA), which authorized the $80 billion boost for tax administration from 2022 to 2031—with $79 billion to the IRS and the remaining funds divided between other Treasury offices and the US Tax Court (table 1). The IRA contained even fewer details than the agenda about how the funding would be used, mandating only the division of the funds among the four broad IRS budget accounts: Taxpayer Services, Enforcement, Operations Support, and Modernization. Over half the funds were allocated to tax enforcement, with just 4 percent set aside for taxpayer services.

TABLE 1

Internal Revenue Service's Mandatory Funding

Fiscal Years 2022 through 2031

	Amounts provided under Inflation Reduction Act of 2022		Recissions: Fiscal Responsibility Act of 2023[a]	Recissions: Further Consolidated Appropriations Act of 2024	Recissions: Total recissions	Mandatory funding after recissions	
	Millions ($)	Percent (%)	Millions ($)	Millions ($)	Millions ($)	Millions ($)	Percent (%)
Taxpayer services	3,181	4.0	-	-	-	3,181	5.5
Enforcement	45,637	57.7	1,400	20,200	21,600	24,037	41.8
Operations support	25,326	32.0	-	-	-	25,326	44.1
Modernization	4,751	6.0	-	-	-	4,751	8.3
Free-file study	150	0.2	-	-	-	150	0.3
Total IRS funding	**79,046**	**100.0**	**1,400**	**20,200**	**21,600**	**57,446**	**100.0**
Addendum: Funding for organizations other than IRS							
Treasury Inspector General for Tax Administration	403		-	-	-	403	
Treasury Department Office of Tax Policy	105		-	-	-	105	
Treasury departmental offices	50		-	-	-	50	
US Tax Court	153		-	-	-	153	
Total mandatory funding	**79,756**		**1,400**	**20,200**	**21,600**	**58,156**	

Source: Inflation Reduction Act of 2022, Pub. L. No. 117-169; Fiscal Responsibility Act of 2023; Pub. L. No. 118-5; Further Consolidations Appropriations Act of 2024, Pub. L. No. 118-47.

THE INFLATION REDUCTION ACT

Since the IRA's enactment, Congress has chipped away at the funding—both directly and indirectly. As part of the agreement between Congress and the President to lift the debt ceiling in 2023, the IRA funding was immediately cut by $1.4 billion in FY 2023, with the reductions to be allocated between the enforcement and operations support accounts. (Ultimately, the IRS reduced the enforcement account by $1.4 billion.) Another $20.2 billion was rescinded in the Further Consolidated Appropriations Act of 2024—all coming from the enforcement allocation.

Moreover, Congress did not erect guardrails between the IRA 10-year funds and the annual appropriations. For fiscal years 2022 through 2024, the annual IRS appropriations have been frozen at $12.3 billion for taxpayer services, enforcement, and operations support. Appropriations for business systems modernization have been eliminated, with the justification that the IRA funds will be used for technological advancements.

In total, the IRS had spent $5.7 billion—or 10 percent—of the IRA funding as of March 2024. However, $2 billion was used to pay for normal operating expenses because of the shortfalls in the annual appropriations (TIGTA 2024).

STRATEGIC OPERATING PLAN

A week after the IRA's passage, Treasury Secretary Janet Yellen instructed the IRS to release an operating plan within six months.[1] Along with more details on how the funds would be used, Secretary Yellen also requested that the plan include metrics for measuring performance.

2023 STRATEGIC OPERATING PLAN

In April 2023, the IRS released a 150-page strategic plan (IRS 2023b). Center to the IRS's Strategic Operating Plan were five objectives:

1. Dramatically improve services to help taxpayers meet their obligations and receive the tax incentives to which they are eligible.
2. Quickly resolve taxpayer issues when they arise.
3. Focus expanded enforcement on taxpayers with complex tax returns and high-dollar noncompliance to address the tax gap.
4. Deliver cutting-edge technology, data, and analytics to operate more efficiently.
5. Attract, retain, and empower a highly skilled, diverse workforce and develop a culture that is better equipped to deliver results for taxpayers.

For each objective, the IRS listed "indicators of success" (appendix 1). However, those were not always fully specified. In some cases, they were simply restatements of the objectives. For example, the first indicator of success in dramatically improving services was "increasing service levels."

Other indicators revealed features of the plan but still without specific metrics—for instance, a "wider array of digital options to help taxpayers and tax professionals interact with the IRS and have a more seamless customer experience." And still others described a metric for evaluation but did not set a quantitative target. An example is "decreased percentage of returns filed with math errors or errors related to third-party information reported to the IRS."

The Strategic Operating Plan also listed 42 initiatives aimed at helping the IRS achieve each objective. The explanation of each initiative included a description of what success would look like. As with the overall indicators of success, those descriptions varied in degree of specificity (appendix 2).

2024 UPDATE TO STRATEGIC OPERATING PLAN

With the 2024 release of an update to the Strategic Operating Plan (IRS 2024d), the IRS made strides toward defining metrics and setting targets for many of its objectives (appendix 3). The update matched objectives to outcomes and specified priority efforts for 2024 and 2025.

Consider again the first objective identified in the Strategic Operating Plan of dramatically improving services. In the 2024 update, the IRS cited nine desired outcomes that would indicate achievement of improved services (e.g., "When taxpayers call the IRS, they are able to reach an agent in a timely manner and have high levels of satisfaction with the interaction"). And to achieve that outcome, the IRS listed three priority efforts (e.g., an 85 percent rate of answered phone calls on the IRS helpline during the filing season with an average wait time of less than five minutes).

As in the example above, the priority efforts were sometimes defined in quantitative measures. In other cases, the effort was described more generally, especially when quantifying the success of the action is not feasible: for instance, other priority efforts intended to meet the objective of better taxpayer services were described simply as "improve Where's My Refund? tool" or "assess impact of Direct File."

LEGISLATIVE REQUIREMENTS FOR MEASURING THE IRS'S PERFORMANCE

Setting performance metrics is not a new task for the IRS. Other legislation—dating back over 40 years ago—have required that the IRS and other government agencies evaluate their performance in certain areas. Those provisions include the

- Government Performance and Results Act (GPRA) of 1993,[2] as amended by the GRPA Modernization Act of 2010;[3]
- Paperwork Reduction Act (PRA) of 1980,[4] as amended by the Paperwork Reduction Act of 1995;[5] and the
- Improper Payments Information Act (IPIA) of 2002,[6] as amended by the Payment Integrity Information Act of 2019.[7]

Of the three, the GPRA is the most extensive. Under the act, government agencies must produce annual performance metrics for both services and enforcement. The PRA and IPIA establish metrics for evaluating certain aspects of services and enforcement, respectively.[8]

GOVERNMENT PERFORMANCE AND RESULTS ACT

GPRA requires government agencies to set goals, periodically prepare strategic plans, measure programs' recent effectiveness each year, and set targets for the future. If an agency does not meet those goals, it must produce a performance improvement plan.

Under GPRA, the performance metrics must measure or assess a program's *outputs*, *service levels*, and *outcomes*. An output measure is defined as the tabulation, calculation, or recording of an activity or effort, whereas an outcome measure is an assessment of how well the program achieved its goals. In its instructions to agencies, the Office of Management and Budget (OMB) encourages agencies to use outcome measures when feasible and appropriate but also lists a broader range of performance indicators than referred to in GPRA.[9] Those include measures for *inputs* (time or monetary costs) and *efficiency* (the ratio of the inputs to its outputs or outcomes).

In its fiscal year 2025 budget released in February 2024, the IRS identified 25 performance measures (IRS 2024a) categorized by the relevant IRS budget account—taxpayer services, enforcement, operation support, and business modernization (appendix 4). Most metrics focused on the *outputs*—for example, the number of answered phone calls and other services provided by the IRS to the public or the exam rates and other counts of enforcement activities. A few measured the inputs (such as rentable space feet per person) or the efficiency of the activity (the costs of collecting $100).

But only two performance measures came close to gauging the *outcomes*: the percentage of surveyed taxpayers satisfied with the IRS (an indicator of the effectiveness of taxpayer services) and the share of individual taxpayers who are noncompliant two years after a prior bad act (an indicator of the effectiveness of IRS enforcement actions).

The IRS's list of performance metrics has evolved. For example, in response to growing interest in equitable treatment of taxpayers, the IRS added three new output metrics showing the number of newly undertaken audits of high-income taxpayers, partnerships, and large corporations beginning in the fiscal year 2022 budget.[10]

PAPERWORK REDUCTION ACT

Even before GRPA, government agencies were required to report on at least one performance element. Under the Paperwork Reduction Act of 1980 (as amended in 1995), agencies—including the IRS—must annually release estimates of the burdens imposed on individuals and businesses by filling out forms. For the IRS, those forms include (but are not limited to) tax returns, W-2s, and 1099s. The burden estimates are partially based on taxpayer surveys.

The estimates for individual and corporate income tax returns are broken into two components:

- Hours spent on each of the following categories—recordkeeping, tax planning, form completion and submission, and other
- Total out-of-pocket expenditures, ranging from payments to preparers and purchases of tax return preparation software to much smaller items such as copying costs and postage

In combination, those components are typically referred to as taxpayers' compliance burden. While the IRS has devoted substantial resources to developing compliance burden measures, the PRA's mandate is limited and focuses solely on one aspect of taxpayers' interactions with the IRS—the costs of completing each IRS form. Thus, the measures understate the total compliance costs incurred by taxpayers—when, for example, waiting for an IRS operator to answer the phone, responding to a request for documentation to support a claim of a child dependent, or being audited.

IMPROPER PAYMENTS INFORMATION ACT

Although the IRS GPRA performance measures do not include estimates of tax noncompliance, the IRS is required to report the amount of erroneous payments of several tax credits each year. Under the Improper Payments Information Act of 2002, as amended by the Payment Integrity Information Act of 2019, each agency must identify programs and activities that "may be susceptible to significant improper payments." Improper payments are defined as any payment that should not have been made or was made incorrectly (either too much or too little) under the law.[11]

Since the launch of IPIA, the OMB and the Treasury Department have included the earned income tax credit (EITC) in the list of programs subject to improper payment reporting. The list has expanded to include three other tax credits: the additional child tax credit (the refundable portion of the child tax credit), the American Opportunity Tax Credit, and the premium assistance tax credit.[12]

The common feature that differentiates those four credits from other tax provisions is that they are partially or fully refundable, meaning that credit claimants can receive payments even if they do not have any income tax liabilities. To the extent that the credits exceed income tax liabilities, the payments are counted as outlays in the federal budget, which is likely considered a justification for the inclusion of those credits with more conventional spending programs in the improper payment analysis.

But the refundable nature of the four credits, combined with income caps on eligibility, also means that lower- and middle-income families are the segment of the population most likely to receive those benefits. Hence, only noncompliance among those groups is required to be reported annually, even though the estimates of improper payments are derived from some of the data used to measure the tax gap—the National Research Program, a nationally representative sample of all individual income tax returns, selected randomly for audits.

PERFORMANCE MEASURES AND THE IRS MISSION

The current IRS metrics are a patchwork of measures mandated by various legislative or administrative requirements. Considered together, they are neither comprehensive nor cohesive.

The IRS's mission statement might be an appropriate starting point for developing a more holistic set of metrics. That statement says that the IRS's mission is to "provide America's taxpayers top quality service by helping them understand and meet their tax responsibilities and enforce the law with integrity and fairness to all." Outcome measures would focus on taxpayer services, enforcement, and equitable treatment of taxpayers to evaluate the IRS's achievement of its mission statement. Output, input, and efficiency metrics could provide additional context as to how the IRS uses the tools at its disposal to achieve those outcomes.

Regardless of the type of performance measure, two issues must be resolved in its choice, design, and implementation—the purpose of the metric and the baseline against which the activity's performance should be measured.

PURPOSE

From outside the agency, the IRS performance measures may resemble report card grades to evaluate the agency's funding. The most recent example was the response to estimates of returns on investment during the deliberations over the IRA funding. Though Treasury and Congressional Budget Office (CBO) analysts disagreed on the amount of revenue that could result from an increase in the IRS's enforcement budget, they both found that the net yield would be positive. That finding played a significant role in the ultimate passage of the IRA funding.

Less visible is the use of performance metrics to highlight trade-offs between programs. Consider statistics of the number of phone calls answered. The IRS issues press releases about the percentage of phone calls to the agency that are responded to during the filing season, reporters write about those results, and lawmakers query IRS and Treasury officials about those numbers during hearings about the filing season and appropriations.

Of equal importance, however, is that the same people who answer the calls also open the mail, and one task can divert resources from the other. In the aftermath of the COVID-19 pandemic, as the IRS staff tried to work through a backlog of unopened paper tax returns, the trade-off between the two tasks became much more visible—in part due to the comments of National Taxpayer Advocate Erin Collins.[13] Bundling together performance measures enables the IRS and others to recognize the trade-offs between goals and make informed judgments as to which activities to prioritize.

Even less transparent is how the IRS uses performance measures to determine how to fix a program with a less-than-satisfactory performance measure. Some exceptions exist, most notably in the IRS's recent focus on audit rates for Black and White taxpayers. In presentations (such as at the IRS–Tax Policy Center research conferences in 2023 and 2024), IRS researchers (Anderson et al. 2024, Hertz et al. 2023) have not only presented data on racial disparities but have also discussed their findings from a more in-depth analysis of the reasons for those differences and how their research has led to changes in how the IRS administers audits.

BASELINE

Another issue common to all metrics is the baseline for observing changes in performance. Often, changes in performance are measured from one year to another or in some cases, back to a year that supports an argument in favor of increases or cuts in funding for the IRS or other legislative actions. For example, supporters of increases to the IRS budget compared current audit rates to higher levels in 2010 (when funding was relatively high) or the number of answered telephone calls in 2023 to much lower levels during the heights of the pandemic (when there was a substantial increase in callers asking questions about temporary assistance programs and IRS staff were working remotely or out caring for themselves and others).

However, a simple comparison of two measures from different years does not indicate either improvements or deterioration of the IRS's performance alone. Achieving the IRS's three missions is not solely the agency's responsibility. The IRS does not write the tax code, but complex laws increase the burden of filing a return, open vulnerabilities for avoidance and noncompliance, and may lead to inequitable treatment because of the ways complicated laws affect groups differently. Nor does the IRS control its funding. Changes in the economy also affect the IRS's performance. In the 21st century, lawmakers have turned to the IRS multiple times to rapidly deliver lump-sum payments to individuals—including people who typically are not required to file tax returns because their income is very low—to alleviate the economic burdens caused by recessions and the pandemic.

Ideally, the IRS would develop metrics that could distinguish between outcomes attributable to its actions and those outside its power to influence. One comment in the IRS's 2023 Strategic Operating Plan suggests that the agency recognizes that need, stating that an indicator of success would be if the tax gap fell "relative to tax gap without the resources provided by the IRA" (IRS 2023b).

But that is only a partial step toward isolating the effectiveness of the agency in enforcing the tax code because it ignores the role of changes in tax laws and the economy. For a model of how to disaggregate the sources of changes to the tax code, the IRS could turn to the budget forecasts of the OMB and CBO. The agencies break down the differences between actual budget data and their prior projections into three

categories—changes due to (1) revised economic assumptions, (2) technical adjustments, and (3) newly enacted legislation.

The IRS is partway there. In its reports on the tax gap, the IRS compares its results to those from the prior study and then decomposes the differences into two categories: (1) updated methods and (2) other factors. For example, the IRS estimated that the annual voluntary compliance rate over the 2014-16 period was 85 percent—up by 1.4 percentage points from the annual rate over the prior three years, of which only 0.1 percentage point was due to revisions in methodology (IRS 2022b).

Currently, the IRS has two outcome metrics for taxpayer services: the GPRA metric for taxpayer satisfaction and the PRA's measure of compliance burdens. In addition, there are eight GPRA output measures—most prominently, the number of telephone calls answered. Those measures could be expanded to provide more insight into the quality of the IRS's services for taxpayers.

OUTCOME: TAXPAYER SATISFACTION

To measure individual taxpayers' satisfaction, the IRS relies on information collected by the American Customer Satisfaction Index (ACSI 2023), a private company founded by researchers at the University of Michigan. Each year, ACSI releases a report on citizen satisfaction with the federal government with breakouts for cabinet departments. The information in the report is based on interviews with a random sample of individuals, with their responses serving as input into an econometric model that derives scores of citizens' satisfaction ranging from 0 to 100.

The public report typically does not include those metrics for all sub-cabinet agencies. However, the IRS obtains the results for filers and includes the value in the agency's annual performance measures. In fiscal year 2022, the IRS's customer satisfaction score was 69. That aggregate score, however, illustrates one of the challenges of a broad performance metric: In this instance, the metric reveals overall satisfaction, but it is insufficient to identify the IRS's strengths and weaknesses. Consequently, the taxpayer-satisfaction metric informs policymakers of the IRS's overall performance but does not provide any insight into the agencies' weaknesses and areas for improvement.

ACSI collects more in-depth data, which would be useful for program evaluation. For example, the ACSI identifies four drivers of citizen satisfaction with the federal government:

- Efficiency and ease of government processes
- Ease of access and clarity of information
- Courtesy and professionalism of customer service
- Perceptions of government websites

The ACSI reports often include a score for each attribute, but only at the government-wide level. Yet, data for departments and agencies might be more useful in at least pointing in the general direction of the source of dissatisfaction.

While further details would be desirable, methodological concerns may constrain the use of the ACSI data. Relative to government household surveys, the description of ACSI's methodology is sparse on its website. But

one anomaly stands out. The size of the survey fluctuates significantly from year to year: 1,291 in 2020 to 2,126 in 2022. In 2023, the sample size shrunk to 847. Perhaps related to that substantial decline in sample size, the IRS did not report a score for taxpayer satisfaction in 2023 in its fiscal year 2025 budget, citing ongoing updates to the methodology.

Another potential source of information is the Comprehensive Taxpayer Attitude Survey or CTAS (IRS, 2022a). That survey is rooted in the Executive Order 12862, issued in 1993, which required agencies to survey customers—people or entities directly dealing with the organization—regarding their satisfaction with its current services.[14] The survey found that 75 percent of taxpayers reported being very or somewhat satisfied with their personal interactions with the IRS in 2021—about 5 percentage points higher than the ACSI found.

One option would be to replace the ACSI with the CTAS as the source of information on taxpayer satisfaction. The large sample size—2,099 adults—may facilitate reliable analysis of subgroups. And because the survey is solely about the IRS, questions can be tailored to address the concerns of taxpayers. However, most of the questions in the 2021 survey focused more on attitudes about the IRS and tax system and did not provide much insight into the specific administrative challenges faced by taxpayers.

A promising sign is President Biden's 2021 initiative to evaluate certain government services, drawing on techniques used to study user experiences in other sectors. As a consequence of the "Executive Order on Transforming Federal Customer Experience and Service Delivery to Rebuild Trust in Government" (Executive Order 14058),[15] the OMB designated the IRS as one of 38 "high-impact service providers" (HISP) in the federal government.[16] HISPs were selected based on the size of their client base or critical impact on those served—both criteria that apply equally to the IRS. Currently, OMB is working with each HISP to develop and implement users' feedback surveys that will be used to derive scores for seven categories: trust, satisfaction, effectiveness, ease, efficiency, transparency, and the quality of interactions with employees.[17] Implementation of this executive order is still in the nascent stage and, at least for the time being, is limited to prioritized services. For the IRS, the current priorities are taxpayers' experiences with return filing and online accounts.

OUTCOME: COMPLIANCE BURDENS

Since the 1980s, the IRS has produced estimates of compliance burdens, relying in part on random surveys of taxpayers. The first surveys asked respondents to report the amount of time they spent on recordkeeping, learning about the law and form, completing the tax form, and transmitting it to the IRS. The survey data were then matched to respondents' tax returns, and the matched data formed the basis of a mathematical model (the ADL model, so called because the survey and modeling were conducted by the Arthur D. Little company).

The ADL model had several shortcomings, which became more problematic over time as people became increasingly reliant on alternative methods of filing. First, the survey did not ask respondents about their monetary costs, including payments to preparers. Nor did the model anticipate the shift from paper returns to preparation software and electronic filing.

While the ADL model is still used for many forms, the IRS began shifting in 2003 to a new approach for estimating compliance burdens for individual and business income tax returns (IRS 2023a). As with the ADL model, the new approach begins with surveys of random samples of individuals and businesses. But the new surveys ask about out-of-pocket expenses as well as hours. Moreover, the random sample of surveyed individuals is stratified by preparation method and complexity category (ranging, for example, from low complexity for wage and salary income to high complexity for partnership income).[18] For the business survey, companies are divided into groups based on their organizational structure and size. The data are then used to build the individual and business burden models, in which the logarithm of the burden—both the monetarized hours and the out-of-pocket expenses—is linearly related to a set of explanatory variables, including income for individuals and assets and receipts for businesses.

The estimates for individual and corporate income tax returns are broken into two components:

- Hours spent on each of the following categories—recordkeeping, tax planning, form completing and submission, and other
- Total out-of-pocket expenditures, ranging from payments to preparers and purchases of tax return preparation software to much smaller items such as copying costs and postage

For example, the 2023 instructions for the individual income tax return (1040) show that taxpayers, on average, take 13 hours to complete their tax return, with nearly half that time devoted to recordkeeping. In addition, they spend an average of $270. Average costs were higher for filers with business income (24 hours and $560) and lower for other filers (9 hours and $150).[19]

Despite the upgrades, the current measures of compliance burden fall short of measuring the IRS's performance in providing taxpayer services—as well as the burdens of interacting with the IRS during an enforcement action.

First, the reported measures of paperwork burden are not broken down by taxpayers' characteristics—other than those that might be inferred by their completion of a form (e.g., we can infer that the filer who attaches a Schedule EIC to his or her tax return is also reporting labor income, total income below a certain threshold, and probably children). Less can be inferred from knowing the compliance cost of completing a 1040 because the form is used by all types of filers (especially since the elimination of the simpler 1040A and 1040EZ in 2018).

A few IRS studies have provided additional information about the association between the paperwork burden and observable characteristics. Because the sample is stratified by the complexity of the return, IRS and Treasury Department researchers provided more detailed information from the 2010 survey about the incidence of taxpayer burdens by the complexity of tax items on individuals' returns (Marcuss et al. 2013). Over half of the compliance costs incurred in the individual income tax were associated with reporting and substantiating income, even for relatively simple returns. As discussed in the section on equity measures, other studies have constructed distributions of paperwork burdens by income group. More analysis of the burden distribution by the presence of children, age of taxpayer, and race and ethnicity would also be informative.

Second, the compliance burden measures meet the requirements of the Paperwork Reduction Act, but it is far from a comprehensive metric of the burden of interacting with the IRS. For example, the burden models do not include prefiling correspondence and discussions between the IRS and the taxpayers (or their advisers) to obtain guidance, such as a letter ruling, nor do the current measures include the burdens attributable to postfiling interactions between the taxpayer and the IRS, though questions about the costs of those interactions were included in at least one of those surveys. An analysis of that data from the 2012 survey indicated that the compliance costs associated with examinations could be as much as $900 more than the costs associated with filing a return for affected taxpayers; however, because only a small share of taxpayers deal with the IRS after filing a return, nearly 60 percent of aggregate compliance costs were incurred before tax returns were filed (Guyton and Hodge 2013). Regularly updating this type of analysis would provide a fuller picture of the costs entailed with dealings with the IRS.

Both preceding challenges have been long recognized by the IRS. In 1998, an IRS study group concluded that the ideal burden estimation model would provide compliance costs by type of tax, taxpayer, and activity (GAO 2000). The group also recommended that burden measures account for all prefiling, filing, and post-filing activities (including enforcement activities).

Third, the compliance burden surveys are restricted to people who actually file tax returns. Yet, the complexity of the tax system potentially burdens people who do not file a tax return because they are not required to do so. Many would have received a refund of over-withheld taxes or a refundable tax credit if they had filed. Those nonfilers may have incurred compliance costs, especially if they had tried and failed to navigate the tax filing process.

Finally, the term "compliance burden" is a misnomer. As measured, the costs include those borne by people not complying with the tax code and by taxpayers following the law. The current measures of compliance costs may be pushed up by taxpayers who search for strategies to avoid or evade their tax liabilities. The average costs could also be deflated by taxpayers who do not read or understand the instructions and thus make inadvertent errors.

Distinguishing between the costs incurred by compliant and noncompliant taxpayers would be especially useful if the scope of the compliance burden studies was permanently expanded to include post-filing activities. The costs involved in an audit can be viewed as part of the penalty when the affected taxpayer is, in fact, noncompliant, but they are unambiguously a burden when the compliant taxpayer must undergo the pain of an unnecessary examination. Linking the measure of compliance burdens to noncompliance research, if possible, would enable the IRS to distinguish between costs incurred by compliant and noncompliant taxpayers.

OUTPUT: TELEPHONE SERVICE

Among the most cited IRS performance measures is the share of telephone calls to the IRS that are answered—the level of service (LOS). It may also be one of the most misunderstood measures.

Consider, for example, reporting of telephone service during 2023—the first filing season after the enactment of IRA. At the close of the 2023 filing season, IRS Commissioner Werfel heralded the historically large increase in the LOS from the prior year—from 16 percent in 2022 to 85 percent in 2023 (Werfel 2023)—the target set by Treasury Secretary Janet Yellen shortly after the passage of IRA.[20] An IRS announcement in April 2024, however, stated that the LOS during the 2023 filing season had been just 84 percent—sparking a reporter's investigation as to whether the IRS had actually missed by a percentage point the 85 percent target set by Yellen for that year (Rifaat 2024).

Other estimates of answered phones in 2023 differed by much more than a percentage point. The Treasury Inspector General for Tax Administration reported that only 52 percent of calls had been answered through May 2023—up from 29 percent in 2022 over the same period (TIGTA 2023). And even estimates by the IRS can differ by much more than a percentage point. In its congressional justification for the proposed fiscal year 2025 budget, the IRS showed that phone service increased from 17 percent in 2022 to just 52 percent in 2023 (IRS 2024b).

One reason for those differences is timing. In response to the reporter's query about the one-percentage-point difference for 2023, a spokesperson for the Treasury Department speculated that there might have been a slight data lag between "Tax Day itself and the end of the filing season." (Rifaat 2024). [21]

The much greater gap between the numbers in the April press release and the congressional justification the following year is due to the former covering just the filing season (from January to April) and the latter representing the entire fiscal year (from October 2022 through September2023). The lower estimates for the entire year reflect shifts in priorities for customer service representatives during the year—from being responsive to taxpayers' questions as they prepare their tax returns to later inputting data from paper returns and responding to taxpayers' correspondence.

The specification of the telephone response rate also contributes to different estimates. For many years, the IRS focused solely on the number of attempted toll-free calls routed to Accounts Management—the line for callers seeking general tax information and updates on tax returns, refunds, and balances due. The level of service (LOS) is the percentage of callers who speak to a customer service representative or receive prerecorded informational messages.[22] TIGTA's measure, the level of access, includes calls diverted to targeted automated lines based on the callers' responses to prompts. However, TIGTA also limits its telephone metric to attempted calls during the IRS open hours.

Beginning with its fiscal year 2024 congressional justification, the IRS introduced a new metric for measuring the performance of customer service representatives—the LOS(A). This measure includes callers who received answers to their questions through an automated tool, though the IRS did not also adopt the TIGTA restriction of only including calls made during working hours. Relative to the original LOS, the LOS(A) was 22 percentage points higher in 2022 and 15 percentage points higher in 2023 (IRS 2024b). Currently, both the LOS and LOS(A) are presented in the IRS's congressional justifications, though it is not obvious which measure is now being used in the April releases.

A third potential source of misunderstanding regards the scope of coverage. The IRS's LOS measures (as well as TIGTA's LOA) are limited to calls routed to Accounts Management. While customer service representatives fielded about 18 million calls in 2023, the IRS received more than 50 million calls (IRS 2024a)—including calls to collections, the refund hotline, Taxpayer Protection Service, to establish installment agreements, or by practitioners seeking priority service.

Beyond the measurement issues, a single-minded focus on the number of calls answered does not give a full picture of the quality of telephone service. Another performance measure, based on a sample of calls, considers the accuracy of the information provided by the customer service representative, and the IRS sometimes separately reports on the wait time and the duration of the call in testimony and press releases. But as National Taxpayer Advocate Erin Collins (2023) points out, other helpful metrics are still missing, including the number of times a taxpayer hangs up because of the length of the wait, whether the taxpayer's issue is resolved during the call, and the taxpayer's perception of their experience with the customer service representative.

Nor do the current performance measures shed much light on the trade-offs in choosing the resources to devote to telephone services. As noted above, the lower year-round LOS estimates reflect shifts in customer service representatives' tasks during the year. The phones do not stop ringing (though likely in lower numbers) after Tax Day, but other delayed tasks take precedence when the filing season ends. Some insight into the output associated with those other tasks is provided by a second performance measure—the number of

accounts management and correspondence work to be processed in inventory—but it is difficult to interpret without additional data on the composition of the inventory.

Context matters in other ways. Technological advances—such as more-accessible information on the IRS website and chatbots—may reduce reliance on telephone service. But if that means that a greater share of answered calls involves complicated questions, the LOS might fall as the waiting period and duration of calls lengthened. Changes in the tax code or unanticipated external events (such as a devastating hurricane or pandemic) may also pressure taxpayer services. Comparisons across years that do not account for factors outside the control of the IRS will make the IRS's performance look weaker.

ENFORCEMENT

Eleven of the GPRA performance measures concern the IRS's performance enforcing the tax code, but only one comes close to being an outcome measure—the repeat noncompliance rate. Still, there are three widely cited enforcement metrics—the tax gap (an outcome measure), audit rates (an output metric), and the return on investment (an efficiency measure). At various times, Congress applied restrictions to developing and using the tax gap and return-on-investment metrics. Both have become more visible after the substantive inflation-adjusted cuts to the IRS budget after 2010.

OUTCOME: TAX GAP

Since 1964, the IRS has periodically conducted studies of tax noncompliance. To many, the tax gap—the difference between the taxes owed and the tax paid on time—may be viewed as the ultimate measure of the IRS's performance as an enforcer of the tax code. Yet, the compliance studies are not mandated, and the tax gap is not included in the performance measures. Indeed, Congress denied funding to continue the studies after 1988 because of concerns about the burdens imposed on individuals selected at random for audits. Funding was restored in the early 2000s only after the IRS committed to revamping the studies to reduce the burden on individual taxpayers.

In its most recent study of noncompliance, the IRS estimated that the gross tax gap was $496 billion per year, on average, for tax years 2014 through 2016—or 15 percent of total tax liabilities owed by individuals and businesses (IRS 2022b). Late payments and enforcement revenue reduced the annual tax gap by $68 billion to $428 billion, on net, and the amount of unpaid taxes to 13 percent of the amount owed.

The IRS estimates the tax gap using information from various sources, including a sample of taxpayers selected randomly for audits, operational audits, and household survey data. By far, the most dominant data source is the National Research Program (NRP) audits of individual taxpayers.

The NRP starts with a stratified random sample of individual income tax returns that are selected for audit. The scope of the audits, however, depends on the complexity of the tax return

- For the simplest returns, if the IRS can reconcile reported amounts with information supplied by third parties (e.g., W-2s and 1099s) and there is no indication of any significant compliance issue, the IRS does not follow up with the taxpayer.
- For somewhat more complicated returns, the IRS will conduct correspondence audits that usually focus on just a few items on a return.
- For the most complicated returns, the IRS will conduct a face-to-face interview with the taxpayer at an IRS office, the taxpayer's home, place of business, or accountant's office.

At the end of the audit, the examiner makes a recommendation (additional tax, no change, or a refund).

The IRS's tax gap research reveals important sources of noncompliance and sheds light on the potential amounts of unpaid taxes that could be collected under current law. But while the NRP is generally well-designed, it may overstate some types of noncompliance while understating other types. Overstating may occur due to the NPR's reliance on the examiner's recommendation. After the audit is completed, taxpayers can appeal or take the dispute to court, but a resolution in the taxpayer's favor does not reduce the tax gap estimate.[23] While the IRS researchers can potentially monitor post-audit abatements (though those disputes may be lengthy), the more challenging task would be identifying when compliant taxpayers do not dispute the examiner's recommendation because of a lack of resources or fear of the IRS (Guyton et al. 2024).

One unambiguous omission is unreported income from criminal activities. Noncompliance attributable to illegal-source income is excluded from the tax gap estimates, partly because of the extreme challenges of observing or estimating the gains from crime.

For some types of income, the IRS actively looks for underreporting but may lack sufficient information or resources to detect most of the noncompliance. Income from partnerships and foreign sources is particularly difficult to observe and verify. In both cases, the challenge of verifying income is compounded by difficulties in tracing the income to the owner. That complicated web makes it difficult to trace the partnership income from the entity to the actual partner who is liable for the tax.

To adjust for undetected unreported income, the IRS uses a methodology called detection-controlled estimation (DCE). The DCE is premised on the assumption that the "best" auditors detect the most underreported income and that those best auditors are the ones who recommend the largest upward adjustments in types of personal income, controlling for observable characteristics of the cases assigned to each examiner. But if the "best" examiners are also the ones who are the most aggressive and make questionable recommendations, the tax gap will be overstated.

The DCE adjustments can be substantial. Without the DCE adjustments, annual individual income was estimated to be underreported by $145 billion, on average, over the 2014-2016 period. The DCE adjustments nearly doubled the estimate of underreported income—up to $278 billion. The estimate of underreported income of sole proprietors increased by 135 percent—from $34 billion to $80 billion—because of the DCE adjustments (GAO 2024). According to GAO, the IRS has not conducted a thorough analysis to determine the causes of these substantial adjustments but has embarked on pilots that may provide more insight. But concerns about the DCE methodology take on greater urgency because of reductions in the NRP sample size—from 14,200 individual income tax returns in 2015 to 4,000 in 2021.

Some countries have adopted other approaches to account for unobserved income. Before 2020, the United Kingdom's HM Revenue and Customs (2020) used the IRS's DCEs to correct for underreporting of income in their tax gap estimates. After an evaluation by the International Monetary Fund in 2013, the United Kingdom began investigating ways to develop multipliers that better reflected the British tax system. Because they found they did not have sufficient observations to build a DCE model, they developed an alternative approach that relies on a panel of experts—including experienced examiners—to estimate how much tax would be undetected in hypothetical audits, involving different types of issues, availability of third-party information, and the degree of cooperation from the taxpayer. This approach uses the Delphi technique, in which experts separately assess the hypothetical cases through a series of rounds, to reach a consensus on the multiplier.

A disadvantage of the UK's approach is that the results might not be replicated with a different group of experts. But it also may yield more information about the vulnerabilities in the tax system and how different types of taxpayers exploit those holes. Supplementing the current method with the UK approach might provide the most useful information.

Assuming sufficient resources, the ideal solution would be to improve the IRS's ability to detect errors on taxpayers' returns. Some of the shortfalls could be addressed with more resources. Improvements in detecting partnership and offshore income are already ongoing, as IRS and academic researchers collaborate on the application of artificial intelligence techniques to compliance studies. Information about the final resolution of an audit—at least through appeals—could be tracked using the IRS's Enforcement Revenue Information System (ERIS). ERIS, however, does not follow a case after it enters the judicial system.

The impact of improvements in detection and data on resolutions could result in a higher or lower estimate of the tax gap than the current DCE-adjusted measures. It would likely be a more accurate measure, especially for specific areas of the tax code (such as partnerships) and shift the tax gap studies from being a score of the IRS's overall performance to an evaluation of the compliance vulnerabilities in the tax system.

None of those potential solutions, however, addresses a fundamental challenge in the tax system. The tax gap measures do not fully reflect the complexity of the tax code. There are many gray areas in the tax code, where complexity contributes to different interpretations of what legal avoidance is and what illegal evasion is (Hemel et al. 2022). The gray areas are especially prominent in the tax provisions affecting high-income taxpayers, partnerships, and large businesses—taxpayers who are more likely than others to have the resources to hire tax advisers capable of designing an aggressive strategy open to different interpretations of its legality. A better understanding of the amounts of revenue lost due to aggressive tax avoidance strategies would complement the tax gap and provide a fuller picture of the IRS's ability to enforce the tax code.

OUTPUT: AUDIT RATES

Audit rates took center stage in the debate about the IRS funding. Supporters of increased funding pointed to the overall reduction in audit rates—particularly among high-income taxpayers and large businesses. Others expressed concern about the potential burdens on compliant taxpayers if they were audited (erroneously) after the IRS's funding increased (Knefel 2022). Adding confusion to the discussion of audit rates was that the IRS changed its definition of audit rate as this debate was ongoing.

Before 2019, the IRS defined the audit rate as the ratio of closed audits in a fiscal year to the number of tax returns filed in the prior calendar year. That measure, however, assumed that within a year of filing, audits began and ended. Increasingly, though, the time gap between filing a return and the closure of an audit has extended beyond a year. As such, those audit rate metrics did not measure taxpayers' likelihood of having been audited on their tax return for a specific year.

The IRS introduced a new measure of audit rates in 2019. The revised measure is the share of returns for a given tax year that are ever audited—a cumulative measure that increases over time as more returns from that year are selected for audit. That rate begins to flatten out once the statute of limitations on assessments is past—typically three years after filing.[24] For 2019 tax returns, that point was reached in 2023.

Consider audits of 2019 tax returns filed by taxpayers with $10 million of reported positive income: As of the end of September 2021, just 2 percent of those returns had been selected for audit; by 2023, that share had grown to 11 percent (IRS 2022a, IRS 2024c). The lag reflects the complexity of their returns and the challenges involved in determining which returns to audit in that income group.

Like the previous metrics, the audit rates should be viewed in a broader context. The quality and quantity of audits matters. Some insight into the quality of the audit is the "no-change" rate: A "no change" audit happens when a taxpayer can substantiate their claims of income, deductions, and credits—in theory, a signal that the IRS was not efficient in its selection of that return for audit. The percentage of cases closed is another indicator of efficiency.

Yet, neither a no-change rate nor a percentage of closed cases is sufficient to judge the quality of the audit selection technique. For example, as of 2023, 97 percent of audits of tax year 2019 individual income tax returns had closed (table 2). Of the closed cases, 12 percent resulted in no change.

TABLE 2

Individual Income Tax Returns
Audit Rates, Closed Cases, and No-Change Rates

Tax Year	All Individual Income Tax Returns			$1 Million or More of Positive Income		
	Audit rate (%)	Closure rate (%)	No-change rate (%)	Audit rate (%)	Closure rate (%)	No-change rate (%)
2010	1.0	99.8	15.1	9.1	96.7	37.0
2011	0.9	99.8	12.4	7.2	96.4	37.0
2012	0.8	99.8	13.3	5.5	96.7	31.9
2013	0.6	99.8	10.5	3.5	95.0	21.7
2014	0.6	99.5	9.3	3.1	90.1	21.6
2015	0.6	99.3	9.3	3.0	85.4	24.9
2016	0.5	98.6	10.2	3.1	74.7	24.8
2017	0.5	98.4	11.8	2.5	74.3	25.3
2018	0.3	97.7	11.8	1.8	82.4	30.4
2019	0.3	97.3	12.0	2.1	81.8	36.6

Source: Author's computations derived from the 2023 IRS Data Book for 2013-2019 tax returns; 2022 IRS Data Book for 2012 tax returns; 2021 Data Book for 2011 tax returns; and 2020 IRS Data Book for 2010 tax returns.

High-income tax returns take longer to audit, and the closure rate, as of 2023, was just 82 percent for audits of 2019 returns for taxpayers with income above $1 million with a no-change rate of 37 percent. That high no-change rate may reflect the longer time it takes to begin and complete audits of the type of complicated return that would ultimately have resulted in a change to the taxpayer's return. But even if all the remaining open audits of those very taxpayers led to an adjustment to the taxpayer's tax bill, the no-change rate for their audits would be 30 percent—more than twice the average no-change for all individual filers.

This example illustrates the limitations of performance measures. First, more than one quantitative metric can be necessary to evaluate the performance of a single activity. And second, more extensive research is needed to put the numbers into context. In this instance, do the higher-than-average no-change rates for high-income taxpayers indicate that the IRS's selection tools are inefficient, or does it mean that the wealthy have more resources to challenge the IRS? Conversely, do the lower no-change rates of low- and middle-income taxpayers indicate that the IRS selection tools are efficient or that they are too intimated, busy, or budget-constrained to challenge the agency's assessments (Guyton et al. 2024)? Other types of research methods—such as focus groups or ethnographic studies—may yield information that can place audit rates and no-change rates into context.

EFFICIENCY: RETURNS ON INVESTMENT

In discussions of the IRS's funding, the return on investment has typically been defined as the ratio of the additional tax receipts, interest, and penalties generated by new audit initiatives to the increase in expenditures for those activities. Historically, the ROIs have been limited to the relationship between collections and the salaries of the IRS employees directly involved in the enforcement actions (including examinations, appeals, and collections).

Until recently, the IRS's ROIs garnered little attention beyond the agency and a small circle of budget analysts and officials at the OMB, Treasury, and CBO. For many years, ROI estimates were viewed with some skepticism—partly because data and research were lacking to support the calculations. Another concern was that funding for IRS expansions had sometimes failed to materialize after the first year (as happened in fiscal year 1992 after a five-year expansion had been enacted the prior year) or was diverted to other uses (as happened in fiscal year 1994, when the initiative's funding was used to pay for unfunded but mandated cost-of-living allowances).

The third reason for the lack of focus on the ROIs was the limitation imposed by the Administration's and Congressional guidelines for inclusion of the effects of spending and revenue bills on the federal deficit in official estimates of the cost or savings of legislation. Those guidelines were formalized in the conference report for the Balanced Budget Act of 1997 and are occasionally updated upon agreement by the House and Senate Budget Committees, CBO, and the OMB. Scorekeeping Rule 14 is particularly relevant to the use of ROIs in budget considerations:

No increase in receipts or decrease in direct spending will be scored as a result of provisions of a law that provides direct spending for administrative or program management activities.

According to CBO, Rule 14 was adopted in part to avoid situations where hoped-for but quite uncertain savings are used to offset near-term certain spending increases or revenue decreases in the same legislation (CBO 2014). The rule applies to all direct spending and revenue proposals.

Nonetheless, legislation on budget processes sometimes permitted appropriators to score the revenues from special "program integrity" initiatives, which allowed an increase in enforcement funds for the IRS (and certain other agencies) above the statutory caps on domestic discretionary spending. For those limited

purposes, economists at Treasury and CBO followed a broadly similar methodology for estimating collections from program integrity initiatives.

Both Treasury and CBO start with ROIs provided by the IRS. The IRS derives ROIs from the Enforcement Revenue Information System (ERIS), which was first developed in the 1990s and has expanded since then. The ERIS follows returns through enforcement and collections activities and contains information both on staff hours and the final amounts collected by the IRS. The time it takes to collect the outstanding tax liabilities after the enforcement action and appeal is based on other confidential IRS data. Those collection periods can stretch out over many years.

The IRS-produced ROIs are averages. Research by Holtzblatt and McGuire (2016) described other assumptions used by CBO at that time to transform those averages into marginal ROIs—the amount of revenues attributable to an additional $1 of appropriations—for estimates of IRS program integrity proposals.

- The ROIs would not reach their peak until at least three years after implementation of an initiative because of the time it would take to hire and then train new employees.
- The marginal revenues from an initiative would decline over time as taxpayers discover new ways to avoid or evade tax liabilities at a faster pace than the IRS could develop counterstrategies.
- The marginal revenue from a new program would be smaller than the ROI for an earlier initiative because the IRS would first tackle the "low-lying fruit"—cases where detection and resolution of errors were easiest.
- Only the revenues directly resulting from audits and collections would be included in the estimates. The estimates omitted any additional revenues resulting from an increase in voluntary compliance.

The focus on ROIs intensified with the release of a paper in 2019 by Natasha Sarin and Larry Summers (2019). They presented a multifaceted plan to expand the IRS, estimating that a $100 billion infusion of funds over 10 years would generate $1.1 trillion of additional revenue for a net deficit reduction of $1 trillion. They diverged from the standard CBO methodology by including revenue from increasing information reporting and additional investments in technology and by excluding the impact of any type of audited taxpayers' responses. In the absence of access to internal IRS (such as the ERIS), Sarin and Summers extrapolated from published data and characterized their estimation as "naïve."

The Sarin and Summers's paper laid the groundwork for the IRA funding boost of $80 billion—and sparked a new discussion of how to estimate ROIs especially after Treasury's (2021) and CBO's initial ten-year estimates of the gross revenues raised by the Administration's $80 billion enforcement proposal differed by about one-third ($316 billion and $200 billion, respectively).[25] Surprisingly, CBO's much smaller estimate in 2021 accounted

for two factors that would, on net, increase enforcement revenues but which were excluded from Treasury's estimates: first, an increase—albeit modest—in voluntary compliance and second, the interaction between the funding increase with another proposal (later dropped) to enhance the IRS's ability to detect noncompliance by requiring more information reporting from financial institutions.[26]

Since then, the IRS and Treasury have revamped their methodology for estimating ROIs and revenue effects of IRS funding (IRS 2024c). In combination, the revisions boost the Treasury revenue estimates of IRA's effects by 27 percent.

That revised estimate is somewhat remarkable because the IRS accounted for costs that were omitted in previous estimates. For example, ERIS understates labor costs because it does not include fringe benefits and the time spent by more than one employee on a case at each stage of the enforcement activity—such as supervisors or others who may be brought in to assist or review a case. Nor does ERIS include the fixed costs associated with new hires—rent for additional space, laptops, and so forth.

Some researchers have incorporated those additional costs into their ERIS computations. Research by Holtzblatt and McGuire (2020) added the costs of fringe benefits, while research by Boning and colleagues (2023) also included labor costs of supervisors and other employees who supported the work of those people directly involved in the enforcement activity. The research by Boning and colleagues (2023) also incorporated expenditures attributable to office space and information technology costs as well as expenses incurred by other government agencies. Whereas CBO's estimate of the ROI peaked at about $7, Boning and colleagues determined that the overall ROI would be $2 once all costs are included. For those in the top 10 percent, the pre-voluntary compliance ROI would be $3 for an additional $1.

But the revised IRS's estimates reflect other factors that more than offset the additional costs. First, the revision accounts for improvements in the IRS's ability to detect noncompliance and efficiently allocate workloads. A second adjustment reflects changes in the assumptions about the IRS's productivity over time. The IRS disputes CBO's assumption that ROIs will decline over time—both because of enhancements in audit efficiency and the large backlog of unworked cases.

Finally, the IRS includes improvements in voluntary compliance after audits—though their analysis is far more optimistic than CBO's assumptions in 2021. After reviewing past compliance research, CBO (CBO 2020) concluded voluntary compliance increased overall in response to audits, but that compliance by higher-income individuals—one of the targets of the Administration's plan for enhanced enforcement—did not improve. The IRS's revised estimates are based on newer research by Boning and colleagues (2023) that finds voluntary compliance to rise among taxpayers in all income groups. In particular, they find that the marginal ROI for taxpayers in the top 10 percent of the income distribution increases from $3:$1 to $12:$1 after accounting for

the deterrence effect among audited taxpayers. The authors note that the ROI might be even higher if they also reflected the indirect effects of audits—that is, the extent to which audits deter noncompliance by people who are not audited.

The IRS has also laid out a vision of future extensions of measures of the ROI to incorporate other types of activities, including non-audit enforcement actions, taxpayer services, and modernization of technologies. Because the IRS lacks data that explicitly links the costs of these activities to the resulting revenue, they provide documentation of the success of similar efforts—not limited to tax agencies—in other countries, states, and the private sector and, in some cases, the amount of revenue or cost-savings achieved by those activities. Based on those studies, the IRS estimated the potential savings for two of the initiatives—notices to prompt taxpayers to make estimated payments and improvements in information technologies. Including those projected savings would double their estimates of the revenue from IRA's funding. Those estimates, however, should be viewed as speculative because neither the full scope of the initiatives nor their costs is detailed in the IRS report.

Notably, the discussion regarding the evidence about the potential revenue gains from improvements in taxpayer services is sparse compared with the other potential activities, and the IRS did not provide an estimate of revenues resulting from an enhancement of taxpayer services. An analysis by Mazur and Sarin (2024) suggests that an additional dollar spent on taxpayer services—such as staffing telephone call centers and walk-in taxpayer assistance sites—could yield at least an additional $2 in revenue, though evidence to support that estimate is scarce.

Another direction is to recognize that some IRS's activities—especially in taxpayer services—have intangible benefits that may or may not affect voluntary compliance. Calling the IRS to get confirmation of one's interpretation of a tax law may not result in any change in reported taxes (especially if the taxpayer's initial interpretation was correct)—and thus show zero returns for the monetary costs incurred by the IRS. But a successful interaction between the caller and the customer service agent can generate goodwill for the IRS with perhaps a positive spillover for trust in the federal government.

Finally, while attention lately has mainly focused on ROIs to support additional funding, the metric is also used to inform choices between specific activities. Research by Hodge and colleagues (2015) estimate ROIs for correspondence audits with different targets to determine which maximizes net direct revenues.

The third goal identified in the IRS mission statement is for the agency to enforce the law with integrity and fairness to all. That goal is perhaps the most difficult to monitor. There are many dimensions of equity, but key elements—such as true income (which includes unreported income), race, and ethnicity—are not immediately observable by the IRS. Those barriers to measuring equitable tax treatment are compounded by the challenges detailed above for estimating the overall performance metrics. Several researchers have attempted to overcome those barriers to examine the distribution of compliance burdens, the tax gap, and audit rates.

OUTCOME: DISTRIBUTING COMPLIANCE BURDENS BY INCOME

Researchers at the IRS and the Tax Policy Center have used the IRS's individual compliance model to distribute the compliance burden by income. After monetarizing the time costs incurred by filers, research by Marcuss and colleagues (2013) shows that the average compliance cost as a share of adjusted gross income falls as income rises. Research by Berger and colleagues (2018) shows similar results using a more comprehensive definition of income, but they also find that the average ratio of cost to income is equally high among families in the bottom income quintile and those in the top 95 to 99 percent of the income distribution.

Still, the challenges found in interpreting the aggregate compliance burden become even more problematic when distributing the costs by income. Compliance costs for low-income individuals might be understated because of the lack of data on nonfilers—some of whom might have tried to file to claim a refund but gave up because they did not understand how to fill out a return or how to seek out assistance. Similarly, the burden measures might not capture the costs incurred by filers who begin but do not complete a form or worksheet (for example, to determine if they should pay the alternative minimum tax).

OUTCOME: DISTRIBUTING NONCOMPLIANCE BY INCOME

Analyses related to the distribution of noncompliance have typically been the byproduct of research on the distribution of income. Some researchers have turned to the NRP to fill in gaps on unobserved income—income neither reported on tax forms nor on household surveys—and their findings provide some insight into the distribution of underpayments of taxes. But although the authors of the studies begin with the same data—the NRP—their results differ:

- Research by Johns and Slemrod (2010) found that the percentage of true income not reported to the IRS increases as "true" adjusted gross income grows but peaks among taxpayers in the top 99 to 99.5 percentile.

- Underreporting is relatively constant across most of the distribution of true income but declines among taxpayers with more than $5 million, according to research by DeBacker and colleagues (2020).
- Research by Guyton and colleagues (2023) estimate that underreported income as a fraction of true income rises from about 10 percent in the bottom 90 percent of the income distribution to 16 percent in the top 1 percent where it remains constant or falls.

In large part, those differences result from the authors' treatment of income that is neither reported nor detected by the IRS. While the Johns and Slemrod research distributes DCE-adjusted income, Debacker and colleagues argue that the DCE is, at best, an adjustment to aggregate income and was not designed to correct for underreporting by individual taxpayers. In their preferred analysis, they distribute unreported income prior to the DCE adjustment. Guyton and colleagues start with DCE-adjusted income and add in their estimates of undetected income from partnerships and offshore accounts. Their distributional findings are driven not only by the addition of those two sources of unreported income but also by their assumptions that most of the undetected income from partnerships and offshore accounts is earned by the very highest-income taxpayers. Those assumptions are disputed in the paper by Auten and Splinter (2024).

Perhaps the most telling comment on this research is found in an appendix to the paper by Guyton and colleagues. They demonstrate how the results differ depending on various assumptions about the distribution of undetected income and conclude: "Finally, in light of all the uncertainty here, we can understand why some readers may wish to give up on DCE, at least for distributional analysis" (Guyton et al. 2023, pp. 38). Nonetheless, the authors remain in the camp of distributing DCE-adjusted income.

What are the implications for understanding the distribution of the tax gap? Of the authors, only Johns and Slemrod estimate the distribution of unpaid taxes. They find that although the percentage of unreported income increases as true income grows, unpaid income taxes as a share of the actual tax liability fall as income rises.

In contrast, an analysis by Sarin (2021) begins with the DCE-adjusted unreported income—the result of a sensitivity test described in an appendix to the DeBacker and colleagues paper—and estimates that the top 1 percent of taxpayers are responsible for 28 percent of the tax gap.[27] Notably, though, she computes the distribution of the tax gap by applying the percentages of unreported individual income in each decile to the aggregate tax gap—which also includes noncompliance from not filing an income tax return or underpaying income taxes as well as noncompliance attributed to payroll taxes, corporate income taxes, and estate taxes.

The distribution of noncompliance is an important outcome of a tax system that is equitable in its treatment of taxpayers. But the analysis of this question is deeply intertwined with the methodological design of tax gap studies.

OUTPUT: DISTRIBUTING AUDIT RATES BY RACE

Taxpayers are not asked to state their race or ethnicity on tax returns or other forms supplied to the IRS, and very few individuals interact in person with an IRS employee. Those factors contribute to a perception that the tax system is race-blind, but the lack of data also makes it difficult to determine whether the tax system treats individuals fairly, regardless of their race and ethnicity. To broaden the examination of disparities in the income tax system, researchers are developing methodologies to add race and ethnicity imputations to tax data.

Those methodological breakthroughs enabled researchers at the IRS and Stanford University to investigate racial disparities in the selection of tax audits (Elzayn 2023). They found that Black taxpayers were three to five times more likely to be audited than non-Black taxpayers—largely due to differences in the audit rates of claimants of the earned income tax credit (EITC).

That study also highlights at least two challenges for measuring racial disparities in the tax code. The first is the need to validate new methodologies. The IRS and Stanford researchers relied on a technique called the Bayesian Improved First Name and Surname Geocoding (BIFSG), first developed to analyze racial disparities in health care. BIFSG imputes race and ethnicity based on first and last names and location (in this case, Census block). Research by Derby and colleagues (2024) identified flaws in this approach that can lead to overstating the probability of non-White individuals being identified as White. In the case of the study of audit rates, their finding suggests that the IRS-Stanford study understated the racial disparities in audit selection. Because Derby and colleagues relied on a sample of low-income families that was not nationally representative, their findings are not conclusive.

The study by Derby and colleagues indicates that the development and implementation of race and ethnicity imputations on tax data is still evolving. In June 2024, the IRS announced an agreement with the Census Bureau, which will provide privacy-protected race and ethnicity data to the IRS and could replace the BIFSG method in the future (Anderson 2024).

The second challenge points again to understanding the context of the metrics. Without further analysis, the finding that audit rates are disproportionately higher among Black taxpayers than other filers is open to many interpretations. Since the release of the study, IRS researchers (Anderson et al. 2024; Hertz et al. 2023) have been delving deeper into the audit selection process to identify the sources for the racial disparities in EITC audits and found that contributing factors include incomplete data on eligibility criteria (including the

child's relationship to the taxpayer), concentration of high-risk tax preparers in non-White neighborhoods, and the IRS's emphasis on selecting returns based on overclaimed refundable tax credits rather than understated income taxes.

The analysis of the racial disparities in audit rates demonstrates the important role played by a combination of a performance measure and in-depth analysis. Based on the findings to date, the IRS has taken steps to refine its audit selection methods—including expanding information on children's relationships to claimants and development of a new EITC risk scoring system (Anderson et al. 2024). This is a promising area for future advances in practice based on evidence.

CONCLUSION

The IRS should—and will—be held accountable for the substantial infusion of IRA funds over the next decade. That requires the IRS to measure and report on its progress—especially with respect to meeting the broad goals set by its mission statement for taxpayer services, enforcement, and equitable and fair treatment of taxpayers. However, performance measures outlined in the IRS' 2023 Strategic Operating Plan and a 2024 update are incomplete. Moreover, other performance measures that predate the plan are a patchwork of sometimes unrelated items that were developed in response to legislative mandates—laws that typically applied to all government agencies. Moreover, there are shortcomings in some measures that can lead to misinterpretations by outside observers.

In this paper, I identify shortcomings of nine of the current metrics. Those metrics, however, form the foundation for measures with key refinements that would better inform evaluations of a transformed IRS.

Taxpayer satisfaction. The IRS currently measures taxpayers' overall satisfaction with the agency, based on survey data collected by an outside organization that monitors consumers' attitudes in many different private and public sectors. But this overall measure provides no insight into the sources of taxpayers' satisfaction and dissatisfaction. User experience surveys, with more detailed questions, could yield useful information that would better guide the IRS in its interactions with taxpayers.

Compliance burden. The IRS measures the costs to taxpayers of filling out forms issued by the agency. Those estimates satisfy the government-wide requirements of the Paperwork Reduction Act. But taxpayers interact with the IRS in many ways—for example, waiting for their call to be answered or searching for information about notices on the IRS website—and those other costs are not included in the current measures. Moreover, the name "compliance burden" is a misnomer, because the measures do not distinguish between the costs of compliant and noncompliant taxpayers.

Expanding the survey to cover other types of activities would fill in some of the information about the costs incurred by taxpayers in their interactions with the IRS. Linking tax burden and tax gap data would provide insight into the trade-offs between the IRS's service and enforcement missions, informing the IRS's and lawmakers' decisions as how to allocate funds between the two sets of activities.

Telephone calls. At least twice a year, the IRS releases information on the percentage of calls that are answered by customer service agents. But one figure covers just the filing season, and the other is for the entire year, and they can greatly differ because of changes in service priorities throughout the year. Providing information about the IRS's service performance throughout the year could reveal more about the trade-offs that the IRS makes—between answering the phones and, for example, opening and responding to the mail. Moreover, bundling the data on answer calls with information about the length of the call, the accuracy of the

information, and whether the taxpayer's question was resolved would provide context for quality as well as for quantity.

Tax gap. Examiners cannot observe all underreported income, nor do they typically have the incentives to uncover unclaimed tax benefits. However, the current methodology will overstate the noncompliance if adjustments are based on decisions made by the most aggressive examiners who may err in favor of the IRS. Improving the IRS's ability to detect noncompliance or unclaimed benefits—through trained examiners and technological improvements, such as artificial intelligence—would reduce the agency's reliance on its current statistical methods. The improvements in detection methodologies might increase or decrease the estimates of the tax gap but would potentially provide more insight into the sources and magnitudes of noncompliance.

Audit rates. Audit rates are an oft-cited quantitative measure, but they provide no information on the quality of the activity. Two other measures can supplement audit rates, though rarely receive the same attention: the percentage of cases closed and the no-change rate (the percentage of audits resulting in no change to the tax return). Yet even the bundling of those three statistics is insufficient to evaluate the quality of audits. A high no-change rate may mean that the audit selection is not well-targeted or that taxpayers have the resources to successfully challenge an examiner's finding of underpaid taxes. Accompanying metrics with more in-depth research—possibly ethnographic studies of audited taxpayers—would shed more light on the quality of audits as well as their quantity.

Return on investment. Until the recent debates over the budget shortfalls and funding boosts, little attention was paid to the returns on investments in the IRS. The new interest in ROIs has also focused on the shortcomings in the historic measures: Not all returns to funding are included in the estimates, but neither have all costs. A more comprehensive measure of ROIs could better inform decisions about the level and allocation of the IRS's budget. But singling out the ROI does not acknowledge the nonmonetary returns to investments in the IRS.

Fairness measures. As the discussion above reveals, it is challenging to measure the IRS's effectiveness in meeting its service and enforcement missions. Those challenges are compounded when analyzing the distribution of the services and enforcement metrics. Improving the aggregate measures and considering how to more accurately capture differences by income or racial groups will provide more insight into whether the IRS also meets its mission to treat taxpayers equitably.

Developing or refining mission-related, comprehensive IRS metrics is essential. It will require resources (or diversion of resources from other IRS tasks) when the IRS already faces many challenges to achieving its goals. But improving performance metrics is also an investment. As a first step, IRS could be more transparent when releasing performance metrics—for example, by identifying the omitted data in a metric or by identifying and

discussing the relationships between metrics. Those first steps can build a foundation for more fundamental changes to the way the IRS measures its success and identifies its vulnerabilities, enabling the agency to better achieve its mission goals of providing support to taxpayers, enforcing the tax code, and treating all taxpayers fairly.

APPENDIX 1

TABLE A1

Internal Revenue Service Inflation Reduction Act Strategic Operating Plan: Objectives and Indicators

Objective	Indicators of success
Dramatically improve services to help taxpayers meet their obligations and receive the tax incentives for which they are eligible	• Increased service levels • Increased taxpayer service options and increased access to, and accessibility of, those options • Decreased filing burdens for taxpayers • Wider array of digital options to help taxpayers and tax professionals interact with the IRS and have a more seamless customer experience • Decreased percentage of returns filed with math errors or errors related to third-party information reported to the IRS • Decreased difference between credits and deductions available versus those claimed, as well as fewer inadvertent errors and improper payments • Higher proportion of taxpayers satisfied with the filing process • Increased taxpayer satisfaction with IRS interactions and service • Higher proportion of paperless processes and systems throughout the IRS, from intake to processing
Quickly resolve taxpayer issues when they arise	• Increased notice response rate • Decreased repeat noncompliance rate • Increased share of credits and deductions claimed by those who are eligible • Decreased average days between return filing and contacting taxpayers with issues • Faster resolution of compliance issues after filing • For taxpayers with the ability to pay, increased percentage of new balance-due cases closed within one year of assessment
Focus expanded enforcement on taxpayers with complex tax filings and high-dollar noncompliance to address the tax gap	• Decreased repeat noncompliance rates • Decreased enforcement contact with the IRS for compliant taxpayers • Appropriate compliance activities for priority segments that are sufficient to achieve voluntary compliance • Decreased tax gap relative to the gap without the resources provided by the IRA
Delivering cutting-edge technology, data, and analytics to operate more effectively	• Faster issue resolution for taxpayers as a result of better data storage and access • Improved abilities to detect and address noncompliance • Data and methodology to measure and reduce the gap between taxpayer uptake and taxpayer eligibility for credits and deductions, as well as to reduce inadvertent errors and prevent fraud • Expanded detail and better ability to capture emerging issues in tax gap measurement • Expanded use of analytics to inform tax administration and improve data management and governance throughout the organization • Faster changes in technology, including adding features that help the IRS respond more quickly to legislative and internal changes and deliver new functionality to taxpayers more rapidly • Consolidated total core systems into fewer applications in large domains, advancing the IRS's mission to centralize data and improve access to operational and taxpayer account information

APPENDIX 1

Objective	Indicators of success
Attract, retain, and empower a highly skilled, diverse workforce and develop a culture that is better equipped to deliver results for taxpayers	• Increased proficiency across the workforce in critical skills, such as technology, data analytics, inclusive leadership, and knowledge of tax administration
	• Increased employee retention, engagement, and developmental opportunities
	• Increased number of qualified applicants and quality of hires
	• Increased headcount that reflects future mission staffing requirements
	• Decreased time-to-hire
	• Expanded opportunities for employee developmental opportunities, including increased training, internal and external rotational assignments, coaching and mentoring
	• Workforce composition that more closely reflects the taxpayers serve by the IRS, including an increased demographic representation of historically underrepresented and underserved communities at each level and segment of the IRS
	• Increased agility and flexibility, moving staff to align with business needs

Source: IRS (Internal Revenue Service). 2023. IRS Inflation Reduction Act Strategic Operating Plan. Publication 3744 (Rev. 4-2023). Washington, D.C.: IRS. https://www.irs.gov/pub/irs-pdf/p3744.pdf

APPENDIX 2

TABLE A2

Internal Revenue Service Inflation Reduction Act Strategic Operating Plan:
Initiatives and Measures of Success

Initiative number	Initiative	Measure of success
Objective 1: Dramatically improve services to help taxpayers meet their obligations and receive the tax incentives for which they are eligible		
1.1	Improve the availability and accessibility of customer service	• A larger share of taxpayers will be satisfied with IRS customer service options and perceive that the IRS is "here to help." • More taxpayers will feel they are receiving accurate and timely answers to questions in ways that meet their diverse needs. • Levels of service across all channels will increase, wait times will fall, and compliance with tax obligations will increase. • The accuracy of returns will also rise, including claims for credits and deductions.
1.2	Expand digital services and digitalization	• Customers will find it easier to prepare and submit documents to the IRS. • Processing times will fall for end-to-end digital processes. • Backlogs in paper and manual processes will be eliminated. • The accuracy of translation of return data will increase as the IRS phases in digital (e-filed) and digitalized (paper converted to digital data) receipts. • Digital tools will be accessible for people with disabilities and available in taxpayers' preferred languages. • Expanding digitalization will also lessen the environmental impact of the IRS by reducing paper usage internally and by taxpayers, reducing the need for shipping and minimizing the amount of paper storage required.
1.3	Ensure employees have the right tools	• Taxpayers' satisfaction will rise with the information they receive during customer service interactions. • Employee morale and job satisfaction will increase as they get more of the tools they need to do their jobs. • The workforce will become nimbler, easier to train, and able to deliver service to taxpayers more quickly and effectively.
1.4	Improve self-service options	• The share of taxpayers using self-service options will increase. • Customers will be satisfied with issue resolution times for self-service options. • The need will fall for live customer service as the primary issue resolution tool for customers who use online accounts. • Customer satisfaction with ease of use and availability of online account features will increase.
1.5	Explore direct file	• There will be high satisfaction rate among customers who use the direct file service, if pursued. • There will be continued use of a direct file service by taxpayers from one year to the next.
1.6	Enable taxpayers to access their data	• Customer satisfaction will improve by helping taxpayers access and understand data. • Customer service calls asking for transcripts and other documents that are accessible through the new online tools will decrease. • Taxpayers who use the data to begin their returns will have fewer document-matching issues and make fewer amended returns.

APPENDIX 2

Initiative number	Initiative	Measure of success
1.7	Provide earlier legal certainty	• Voluntary compliance will rise due to increased understanding. • The IRS will receive fewer requests from stakeholders and IRS employees for additional guidance and will be able to focus more compliance resources on issues where there are true legal disputes.
1.8	Deliver proactive alerts	• The accuracy of participating taxpayers' tax reporting and claims for appropriate incentives will increase. • Penalties and interest will decline, and examination cycles will be shorter for taxpayers who receive relevant alerts.
1.9	Help taxpayers understand and claim appropriate credits and deductions	• The share of taxpayers successfully claiming incentives for which they are eligible will increase. • Taxpayers will face fewer barriers to claiming credits and deductions they are eligible to receive as the IRS redesigns policies, processes, and procedures to facilitate access. • There will be fewer inadvertent errors, as well as less fraud and abuse.
1.10	Make payments easy	• The share of timely payments will rise, and accruals of interest and penalties will decline. • Customer satisfaction regarding the ease and transparency of payment processes will increase.
1.11	Build status-tracking tools for taxpayers	• There will be more accurate projections for process and decision timeframes. • Customer satisfaction with the IRS status-tracking tools will increase, including improved satisfaction with accuracy, ease of use, and value of information provided. • The IRS will receive fewer calls related to tracking the status of tax returns, refunds, audits, and other IRS processes.
1.12	Streamline multichannel customer assistance	• There will be higher rates of first-contact resolution for a larger share of issues. • Taxpayers will resolve more issues through self-service and automated channels and require live customer service channels less often. • Customer satisfaction rates will increase related to the ease of communicating with the IRS and of escalating customer service issues through different channels.
Objective 2: Quickly resolve taxpayer issues when they arise		
2.1	Identify issues during filing	• Fewer returns will be rejected, and fewer returns will be sent to error resolution for taxpayers who receive these notifications. • For taxpayers who receive notifications about credits and deductions, claims for appropriate credits and deductions will increase.
2.2	Deliver early and appropriate treatments for issues	• The IRS will help taxpayers resolve compliance issues more quickly. • There will be a higher incidence of correcting or self-correcting issues using less intrusive treatments when appropriate. • Repeat noncompliance will decline, and there will be a higher incidence of taxpayers paying what they owe.
2.3	Develop taxpayer-centric notices	• Response rates and compliance rates will increase among taxpayers who receive improved notices. • Taxpayers will understand what information must be provided to claim credits and deductions. • Taxpayers' rate of using self-service options to resolve issues will likewise increase. • Customer satisfaction measures will increase due to clearer notices and easier response processes.

APPENDIX 2

Initiative number	Initiative	Measure of success
2.4	Expand tax certainty and issue resolution programs	• Participation in tax certainty programs will rise. • Resolution of taxpayers' complex tax issues will be faster. • Post-filing compliance activities for participating taxpayers will decline.
2.5	Offer proactive debt resolution	• Voluntary compliance will increase. • The number of taxpayers who willingly enter arrangements to address their past-due balances in ways that are suited to their circumstances will increase. • There will be reductions in the overall amount of tax debt, the number of taxpayers with past-due balances, and accumulated interest.
2.6	Expand engagement with non-filers	• There will be an increase in non-filers who are detected and contacted in a timely manner so they can voluntarily comply. • There will be an increase in simple options that reflect the reasons why individuals and entities do not file. • The number of taxpayers who self correct (i.e., pay on their own without intervention) will increase. • The number of taxpayers who face additional penalties will fall.
2.7	Use improved data and analytics to tailor timely collections contacts	• There will be a higher percentage of new balance-due cases closed within one year of assessment. • There will be a higher number of communication methods used to contact taxpayers based on their preferred channels, including text, telephone, and email. • Taxpayers will also resolve past-due tax balances more quickly.
Objective 3: Focus expanded enforcement on taxpayers with complex tax filings and high-dollar noncompliance to address the tax gap		
3.1	Employed centralized, analytics-driven, risk-based methods to aid in the selection of compliance cases	• A more agile IRS organization will centralize adoption of enhanced risk analytics to respond more effectively to emerging compliance issues with more appropriate, less burdensome compliance treatments for taxpayers.
3.2	Expand enforcement for large corporations	• The IRS will increase audit coverage and other types of enforcement of large corporate taxpayers to improve voluntary compliance.
3.3	Expand enforcement for large partnerships	• The IRS will increase audit coverage and other types of enforcement of large partnership taxpayers to improve voluntary compliance.
3.4	Expand enforcement for high-income and high-wealth individuals	• The IRS will Increase audit coverage and other types of enforcement of high-income and high-wealth taxpayers to improve voluntary compliance.
3.5	Expand enforcement in areas where audit coverage has declined to levels that erode voluntary compliance	• The IRS will increase enforcement and compliance coverage across all areas where coverage has been too low, including excise tax, employment tax, and estate and gift tax.

APPENDIX 2

Initiative number	Initiative	Measure of success
3.6	Pursue appropriate enforcement for complex, high-risk, and emerging issues	• The IRS will increase enforcement of known, high-risk issues and conduct more timely detection of and response to emerging issues.
3.7	Promote fairness in enforcement activities	• Enforcement actions that appropriately reflect risk and level of noncompliance and address enforcement disparities will be adopted. • Trust in the IRS and voluntary compliance will increase as taxpayers gain confidence that the tax system is administered in a fair manner.
Objective 4: Delivering cutting-edge technology, data, and analytics to operate more effectively		
4.1	Transform core account data and processing	• Volume of programming code in outdated languages will decline as IRS transforms the code to current languages. • A consolidated enterprise program will mitigate the need to build custom applications for specific functions. • More data will be available for taxpayers, analytics, and system-to-system integrations. • Data and online engagement options will improve taxpayers' and employees' experiences.
4.2	Accelerate technology delivery	• There will be robust application of the product-and-platform operating model, with effective partnership between the IT team and the rest of the IRS organization. • Better products, tools, and improvements will be delivered more quickly.
4.3	Improve technology operations	• Modern, flexible technology processes and platforms will be created that will allow the IRS to serve taxpayers better with ongoing data security, a more productive workforce, faster processing times with fewer errors, and faster adaptation to future needs.
4.4	Continue to ensure data security	• The continued evolution of protections afforded to taxpayer data will be ensured. • The IRS will remain one step ahead of increasingly sophisticated cyberattacks. Leading-edge technology will ensure that every action taken or system accessed is appropriate and authorized. • All taxpayer data will be internally encrypted and segmented to limit exposure to threats and compliant with all federal standards and guidelines.
4.5	Maximize data utility	• The IRS will have high-quality, complete data for strategic decision making. • Data will be available for analytics to help clarify taxpayer needs, better identify emerging compliance issues, increase operational effectiveness and innovation, and ensure fair and equitable treatment of taxpayers. • IRS employees will be able to access appropriate data in a usable format to fulfill their responsibilities and serve taxpayers. • Enterprise data platforms will expand taxpayers' access to their own data.
4.6	Apply enhanced analytics capabilities to improve tax administration	• There will be more tailored and effective interactions with taxpayers across IRS programs. • Employees will be better equipped to understand taxpayer problems and provide appropriate and comprehensive service. • Compliant taxpayers will be less likely to be selected for compliance treatments, reducing their burdens. • Taxpayers with compliance issues will benefit from more expedient resolution of their cases as employees will be better equipped with all information relevant to their case.

APPENDIX 2

Initiative number	Initiative	Measure of success
4.7	Strategically use data to improve tax administration	• Better informed, data-driven decision-making, and planning will allow the IRS to deploy staff more effectively where they are most needed. • Innovation in customer service will increase taxpayers' responsiveness. • Data-driven resource allocation and strategic planning will enhance organizational effectiveness and performance.
4.8	Partner to expand insights	• The IRS and other government agencies will benefit from the ability to collaborate while providing state of-the-art data protections. • The IRS will gain new insights that inform policy and program priorities to improve constituent services. • The IRS will benefit from the development of innovative ways to address tax administration challenges, enhanced rigor in validating methods and findings, and continuous learning.
Objective 5: Attract, retain, and empower a highly skilled, diverse workforce and develop a culture that is better equipped to deliver results for taxpayers		
5.1	Redesign hiring and outboarding	• The IRS will improve the end-to-end hiring process and reduce the time-to-hire, regardless of the type of vacancy. • All hiring actions will proceed more smoothly and quickly due to process improvements that maximize efficiency while retaining meritocratic principles. • As IRS improves ability to identify candidates with the skills required to fulfil the duties of a position, newly hired employees will experience a measurable increase in satisfaction.
5.2	Attract a talented and diverse workforce	• The number of qualified, diverse applicants will increase in response to external job searches in an expanded talent pipeline. • The IRS will convert a larger share of applicants to employees as they refine and improve recruiting efforts. • Candidates and hires will better reflect the diversity of the United States as IRS reachs out to historically underserved and underrepresented communities with compelling, accessible opportunities.
5.3	Improve the employee experience	• Employee satisfaction and retention rates will increase. • Employees' perception of workforce flexibilities, work-life balance, workplace inclusion, and leadership culture will improve. • All offices will have needed supplies, and facilities will meet standards to foster a professional atmosphere.
5.4	Help employees grow and develop	• There will be a marked increase in participation in and satisfaction with training, rotational, and learning programs. • The number of offerings and how often IRS update them will likewise increase. • The IRS will retain more people in high-priority, high-turnover populations. • Improving managerial jobs will boost the number of applications for manager roles and increase satisfaction rates among managers. • Employee proficiency in core knowledge and skills—including basic data analysis, problem-solving, cultural competency, leadership, and the fundamentals of tax administration—will rise, allowing employees to take end-to-end ownership of issue resolution. • All employees in leadership roles will participate in a mentorship program, and mentorship scores should rise in FEVS and other employee surveys.
5.5	Develop a data-savvy workforce	• The IRS will be a more data-driven and data-minded organization, where employees at all levels have opportunities to grow their data literacy and apply their data knowledge. • Data and analytics capabilities will support efforts to enhance the taxpayer experience, reduce taxpayer burdens, and improve mission outcomes.

APPENDIX 2

Initiative number	Initiative	Measure of success
5.6	Elevate workforce planning strategy	• The IRS will use talent analytics to improve human capital operations and outcomes at all stages of the employee lifecycle. • Robust forecasting will facilitate more predictable and prioritized hiring, predictable attrition rates, and a prioritized, demand-based annual hiring plan. • Putting analytics to use beyond forecasting will yield a range of human capital enhancements, including attrition reduction and higher performance among new hires.
5.7	Improve organizational structures and governance	• The IRS will adopt a more unified, simpler structure that enables employees to serve taxpayers and tax professionals better and to provide them with an improved experience. • Taxpayers will benefit from improvements in IRS internal operations, especially when they need to address complicated issues that require coordination across IRS teams.
5.8	Build a culture of service and continuous improvement	• The IRS will adopt and implement the right customer-service standards, governance, organizational structure, risk policies, decision rights, leadership performance commitments, and process improvement policies to support cultural change. • As the culture shifts, trust in leadership will increase, driving a corresponding increase in employee satisfaction and feelings of empowerment. • Taxpayer satisfaction regarding issue resolution will increase. • Data gathered over time as cultural baselines are set and monitored will show shifts in culture toward IRS goals. • Processes throughout the IRS will improve, reducing taxpayers' and employees' burdens and yielding gains in efficiency and effectiveness.

Source: IRS (Internal Revenue Service). 2023. IRS Inflation Reduction Act Strategic Operating Plan. Publication 3744 (Rev. 4-2023). Washington, D.C.: IRS. https://www.irs.gov/pub/irs-pdf/p3744.pdf

APPENDIX 3

TABLE A3

Internal Revenue Service Inflation Reduction Act Strategic Operating Plan 2024 Update:
Objectives and Indicators of Success

Objective number	Objective	Outcome	2024 priority effort
1	Dramatically improve services to help taxpayers meet their obligations and receive the tax incentives for which they are eligible	• When taxpayers call the IRS, they are able to reach an agent in a timely manner and have high levels of satisfaction with the interaction.	• Maintain 85% Level of Service on the main taxpayer helpline during the filing season (achieved 88% Level of Service) • Achieve average call wait time of less than 5 minutes on the main taxpayer helpline during filing season (achieved 3-minute average wait time) • Provide a callback feature to 95 percent of callers on priority phone lines to reduce wait time
		• Improvements are made in call centers, while paper processing speeds are kept at high levels.	• Reduce backlog of individual amended returns by 5 percent versus prior year
		• Taxpayers can schedule a timely appointment for face-to-face support across all geographies.	• Increase staffing at more than 350 Taxpayer Assistance Centers across the nation to provide an additional 8,000 hours of in-person support (provided an additional 11,000 hours)
		• Taxpayers in underserved communities will have increased access to IRS services.	• Selected VITA sites will offer free tax return preparation assistance to taxpayers who participate in the "gig" or small business economies • Increase the number of returns prepared by Volunteer Income Tax Assistance (VITA) and Tax Counseling for the Elderly (TCE) sites by 200,000 • Execute Direct File pilot, ensuring access for people with disabilities, Spanish speakers, and taxpayers using mobile devices
		• Individual taxpayers can perform most of their required interactions with the IRS through their Individual Online Account.	• Execute Direct File pilot • Assess impact of Direct File pilot • Expand capabilities of Individual Online Account • Improve Where's My Refund? tool
		• Businesses can perform most of their required interactions with the IRS through their Business Online Account.	• Expand capabilities of the Business Tax Account
		• Tax professionals can perform most of their required interactions with the IRS through their Tax Professional Online Account.	• Expand capabilities of the Tax ProAccount
		• Taxpayers have easy-to-use digital channels for receiving information from and submitting information to the IRS in line with commercial expectations (e.g., banks, insurers).	• Enable taxpayers to digitally submit virtually all correspondence, non-tax forms, and responses to a notice they receive from the IRS • Enable taxpayers to e-file over 10 new tax forms, including but not limited to: 4626, 7203, 7205, 7207, 7208, 7210, 8332, 8936, 8985, and 8986 • Transform twenty priority non-tax forms to have modern, mobile-friendly formats that makes them easy to complete or file electronically
		• Paper that taxpayers choose to send to IRS is digitized at the point of receipt.	• Replace aged scanning equipment and automated mail sorters in high volume locations
		• Once digitized, information is processed throughout the IRS in a digitally optimized manner.	• Digitally transform a small set of priority IRS internal processes (5-10), including but not limited to digital intake, inventory management, and data redaction
		• IRS employees have the hardware, software, and supplies needed to deliver effective and efficient results to taxpayers.	• Replace or update aging individual tools (e.g., laptops, phones, office supplies) • Replace aging capital equipment (e.g., mail sorters, bulk scanners) • Develop and integrate high priority software tools into operations (e.g., advanced tax calculators, modern case management software)

APPENDIX 3

Objective number	Objective	Outcome	2024 priority effort
2	Quickly resolve taxpayer issues when they arise	• Notices are easy to understand and meet taxpayer needs.	• Redesign at least 20 high volume notices
		• IRS Business Units and Office of Taxpayer Correspondence can quickly customize content and recipients of notices via improved technical solutions and streamlined clearance processes.	• Increase the number of notices viewable for individual taxpayers in their online account
		• IRS uses advanced analytics and government and external partner expertise to proactively identify and rapidly minimize the impact of tax scams.	• Coordinate with partners, such as tax software providers, to protect taxpayers by identifying scams and victims at point of filing • Use advanced analytics to measure success and find areas of vulnerability • Improve victim assistance for those that fall victim to tax scams
3	Focus expanded enforcement on taxpayers with complex tax filings and high-dollar noncompliance to address the tax gap	• IRS expands enforcement for priority taxpayer segments by increasing total trained staff and making compliance work more efficient with new tools and processes.	• Expand enforcement presence through hiring in critical staffing areas such as revenue agents, revenue officers, and tax specialists, including those dedicated to high-income earners and large and complex partnerships • Stand up dedicated team to focus on structure of new organization to work complex partnerships and pass through entities • Expand efforts involving digital assets, including work through the John Doe summons effort and the release of proposed regulations on broker reporting • Continue examinations of 76 of the largest partnerships in the United States that represent a cross section of industries including hedge funds, real estate investment partnerships, publicly traded partnerships, large law firms, and other industries • Continue to explore alternative interventions including "soft" notices such as compliance alerts sent to large partnerships with balance sheet discrepancies
		• Taxpayers receive improved fairness in enforcement through new dedicated programs and improved analytics.	• Establish a team and develop workplan to evaluate key enforcement programs for fairness • Invest IRA resources into research that can help identify any disparities across dimensions of race, ethnicity, age, gender, and geography, and refine approaches to compliance and enforcement
		• IRS uses a centralized, data-backed case selection and compliance planning function to more effectively identify and act upon high-priority compliance cases.	• Develop data and research approach to inform and continuously refine compliance strategy needed to promote voluntary compliance • Establish at least one minimum viable product for advanced case selection of priority taxpayer segment such as high income-high wealth

APPENDIX 3

Objective number	Objective	Outcome	2024 priority effort
4	Delivering cutting-edge technology, data, and analytics to operate more effectively	• IRS foundational IT systems are modern, scalable, and easy to maintain to meet the IRS mission.	• Initiate modern individual tax processing to be run in parallel with the legacy Individual Master File (IMF) • Continue to develop an enterprise data platform for use across taxpayer service and compliance functions with easily consumable services • Replace high-volume manual processes through the delivery of "bots" built using robotic process automation (RPA) • Expand the functionality of the new free Information Returns filing portal (Information Returns Intake System – IRIS)
		• IRS leverages technology to improve implementation of the Clean Energy Inflation Reduction Act and CHIPS legislation and Science Act.	• Partner with the Department of Energy (DOE) to enhance and develop their clean energy credit online tools; develop statistics of credit uptake; and ensure data and analytics are available for fraud detection and compliance
		• IRS uses an operating model that delivers technology products in a timely, transparent, and collaborative manner to meet business and operational needs.	• Implement a new operating model, with increased partnership between the IT team and the rest of the IRS organization, allowing the IRS to more quickly deliver better products, tools, and improvements
		• Technology meets all required security standards to ensure data are secure.	• Enhance protection of IRS data through cybersecurity enhancements including multifactor authentication and data-at-rest encryption
		• Data are available for operations and innovation.	• Ensure critical data are available for responsible use of advanced analytics and innovation to improve enforcement, taxpayer services, and operations • Evaluate future access models for responsible use of third party data and complete pilot integration of corporate registry data into IRS data platforms • Create synthetic data to share with external partners to improve insights
5	Attract, retain, and empower a highly skilled, diverse workforce and develop a culture that is better equipped to deliver results for taxpayers	• Employees are enabled with the skills and tools needed for career development and growth.	• Expand training provided to managers to increase their ability to support their employees and the IRS mission • Launch IRS University (IRSU) proof of concepts and finalize the IRSU organizational structure to improve IRS employees' access to developed curricula, "just-in-time" courses, continuous learning, and professional development opportunities
		• The IRS's workforce will be expanded using streamlined, efficient methods for workforce planning, recruiting, hiring, and onboarding quality applicants that represent the American taxpayers.	• Mature IRS workforce planning capabilities by enhancing hiring plans to include additional position requirements to perform critical work • Continue to leverage direct hire authority (e.g. direct hire events) to increase scale and speed of hiring for enforcement positions, including those responsible for high-income taxpayers, large corporations, and complex partnerships • Develop an IRS Employer Value Proposition to highlight how employees can grow their careers, serve the public, and thrive in their personal life • Expedite the average time to hire and support the agency's hiring targets in critical staffing areas such as taxpayer service, enforcement, data analytics, and IRS operations • Establish innovative approaches to reach and develop new pipelines of talent in underserved and underrepresented communities
		• A culture and workplace environment will provide an excellent employee experience, empower employees, and foster diversity and inclusion.	• Update onboarding and orientation pilot programs • Launch Health of the Workforce dashboard, to monitor workforce trends at different levels and teams within the IRS • Reduce volume of Reasonable Accommodation Request Inventory that has exceeded internal processing targets by 50 percent

Source: IRS (Internal Revenue Service). 2023. IRS Inflation Reduction Act Strategic Operating Plan. Publication 3744 (Rev. 4-2023). Washington, D.C.: IRS. https://www.irs.gov/pub/irs-pdf/p3744.pdf

APPENDIX 4

TABLE A4

IRS Performance Measures in Fiscal Year 2025 Budget

Measure	Description	Type of metric
Taxpayer services		
Customer service representative level of service	The number of toll-free callers that either speak to a Customer Service Representative or receive informational messages divided by the total number of attempted calls.	Output
Total ending inventory	The total number of accounts management and correspondence work to be processed in inventory.	Output
Percent of closures to receipts	The number of adjustment cases closed compared to the numbers received.	Output
Level of service (A)	The relative success rate of taxpayers that call seeking assistance from IRS and receive a response to their inquiry by an assistor or through automated responses divided by the total number of attempted calls.	Output
Customer accuracy – tax law (phones)	The percentage of correct answers given by a live assistor on toll-free tax law inquiries.	Output
Customer accuracy – accounts (phones)	The percentage of correct answers given by a live assistor on toll-free account inquiries.	Output
Timeliness of critical individual filing season tax products to the public	The percentage of Critical Individual Filing Season (CIFS) tax products available to the public seven calendar days before the official IRS start of the (individual) filing season. CIFS tax products are those tax forms, schedules, instructions, and publications required by a large number of filers to prepare a complete and reasonably accurate individual income tax return.	Output
Timeliness of critical tax-exempt/government entities and business tax products to the public	The percentage of Critical Tax Exempt/Government Entities (TE/GE) and Business (CTB) tax products available to the public seven calendar days before the official IRS start of the (individual) filing season. CTB tax products are forms, schedules, instructions, and publications required by a large number of TE/GE and Business filers to prepare a complete and reasonably accurate return or form by the filing data occurring during the fiscal year (e.g., income tax, excise tax, exempt organization return, etc.).	Output
Enterprise self-assistance participation rate	The percentage of taxpayer assistance requests resolved using self-assisted automated services.	Output
Taxpayers satisfied with the IRS	The score of taxpayers satisfied with the IRS according to the American Customer Satisfaction Index (ACSI). The All Individual Tax Filer score is calculated from separate ACSI Individual Paper and Electronic Filer Customer Satisfaction Index Scores. Based on a 100-point score.	Outcome

APPENDIX 4

Measure	Description	Type of metric
Enforcement		
Examination efficiency – individual (1040)	The sum of all individual 1040 returns closed by Small Business/Self-Employment, Wages and Investment, Tax-Exempt and Government Entities, and Large Business and International (field and correspondence examination programs) divided by the total number of full-time equivalents expended in relation to those individual returns.	Output
Time to start compliance resolution	The percentage of all individual income tax enforcement cases started within six months of the return posting date.	Output
Time to resolve compliance issue after filing	The median time it takes to close all individual income tax enforcement cases in days (excluding disaster, bankruptcy, and TEFRA cases for exam and collection cases that are not closed as full paid), starting from filing date.	Input
Repeat noncompliance rate	The percentage of individual taxpayers in a fiscal year of individuals with additional noncompliance two years after the initial tax year that contains a filing, payment, or reporting compliance issue, compared to total taxpayers.	Outcome
Collection coverage	The volume of collection work disposed compared to the volume of collection work available.	Output
Exam starts – high-income individuals	The number of examinations started during the fiscal year of individuals with a total positive income of $10 million and above. Total positive income is the sum of all positive amounts shown for the various sources of income reported on the individual income tax return and thus excludes losses.	Output
Exam starts – partnerships	The number of partnership examinations started during the fiscal year.	Output
Exam starts – large corporations (assets greater or equal to $250 million)	The number of examinations started during the fiscal year of large corporate returns reporting assets of $250 million and above.	Output
Costs to collect $100	The cost of collecting $100 is computed as total operating costs divided by gross collection multiplied by 100. Operating costs are comprised of items charged to discretionary appropriations, mandatory appropriations, and user fees.	Efficiency
Criminal investigations completed	The total number of subject criminal investigations completed during the fiscal year, including those that resulted in prosecution recommendations to the Department of Justice as well as those dismissed due to a lack of prosecution potential.	Output
Conviction rate	The percent of adjudicated criminal cases that result in convictions.	Output

APPENDIX 4

Measure	Description	Type of metric
Technology and operations support		
Rentable square feet per person	The amount of rentable square feet the IRS maintains per person requiring space.	Input
Percent of aged hardware	The percentage of all information technology hardware in operation that is past its useful life.	Input
Business systems modernization		
Percent of major IT investments within 10 percent cost variance at the investment level	Number of major IT investments within plus or minus 10 percent variance between planned total cost and projected/actual cost within a fiscal year divided by the total number of major IT investments in that fiscal year.	Input
Percent of major IT investments within 10 percent schedule variance at the investment level	Number of major IT investments within plus or minus 10 percent variance between planned total days and projected/actual days within a fiscal year divided by the total number of major IT investments in that fiscal year.	Input

Source: IRS (Internal Revenue Service). 2024. Fiscal Year 2025 Congressional Budget Justification and Annual Performance Report and Plan. Publication 4450 (Rev. 2-2024). Washington, DC: IRS. https://www.irs.gov/pub/irs-pdf/p4450.pdf

NOTES

[1] US Secretary of the Treasury Janet Yellen, memorandum for IRS Commissioner Charles Rettig, August 17, 2022, https://www.taxnotes.com/tax-notes-today-federal/tax-system-administration/yellen-requests-irs-plan-resource-implementation/2022/08/18/.

[2] Government Performance and Results Act of 1993, Pub. L. No. 103-62, 107 Stat. 285 (1993).

[3] GRPA Modernization Act of 2010, Pub. L. 111-352, 124 Stat. 3866 (2011).

[4] Payment Reduction Act of 1980, Pub. L. 96-115, 94 Stat. 2812 (1980).

[5] Paperwork Reduction Act of 1995, Pub. L. 104-13, 109 Stat. 163 (1995).

[6] Improper Payments Information Act of 2002, Pub. L. 107-300, 116 Stat. 2350 (2002).

[7] Payment Integrity Information Act of 2019, Pub. L. 116-117, 134 Stat. 113 (2020).

[8] Other legislative acts have included provisions concerning performance metrics. The Internal Revenue Service Restructuring and Reform Act of 1998 (Pub. L. 105-206, 112 Stat. 685) required the IRS to set performance goals for organizational units and the establishment of a balanced performance measurement system. The IRS subsequently established a system (Establishment of a Balanced Measurement Program, 64 Fed. Reg. 42835 [August 6, 1999]) composed of three elements: customer satisfaction measures, employee satisfaction measures, and business results. Although the metrics were developed for monitoring individual units within the IRS, they are used—when appropriate—in establishing metrics for the entire agency under the Government Performance and Results Act. The Taxpayer First Act (Pub, L. 116-25, 133 Stat. 981 [2019]), mandates that the IRS identify metrics and benchmarks for quantitatively measuring the progress of the IRS in implementing a comprehensive customer service strategy.

[9] Executive Office of the President, Office of Management and Budget, Preparation, Submission, and Execution of the Budget, Circular No. A-11 (2017).

[10] High-income individuals are those with a total positive income of $10 million and above. Total positive amounts shown for the various sources of income reported on the individual income tax return and thus excludes losses. Large corporate returns are those reporting assets of $250 million and more.

[11] In this context, the term "significant" means that, in the preceding fiscal year, the sum of a program or activity's improper payments and payments whose propriety cannot be determined by the executive agency due to lacking or insufficient documentation may have exceeded (1) $10,000,000 of all reported program or activity payments of the executive agency made during that fiscal year and 1.5 percent of program outlays or (2) $100,000,000.

[12] "Payment Accuracy," US Federal Government, accessed July 17, 2024, https://www.paymentaccuracy.gov/.

[13] Michelle Singletary. "This Tax Season, the IRS Answered Just 10 Percent of Taxpayer Calls," *The Washington Post*, June 24, 2022, https://www.washingtonpost.com/business/2022/06/24/irs-taxpayer-calls/.

[14] Exec. Order No. 12862, 58 FR 48257. https://www.google.com/books/edition/United_States_Code_Congressional_and_Adm/H9QkAAAAMAAJ?hl=en&gbpv=1&dq=Exec.+Order+No.+12862,+58+FR+48257&pg=SL2-PA73&printsec=frontcover

[15] "Executive Order on Transforming the Federal Customer Service Delivery to Rebuild Trust in Government." The White House, December 13, 2021, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/12/13/executive-order-on-transforming-federal-customer-experience-and-service-delivery-to-rebuild-trust-in-government/.

[16] "Federal Customer Experience," US Federal Government, accessed on July 17, 2024, https://www.performance.gov/cx/.

[17] Executive Office of the President, Office of Management and Budget, Preparation, Submission, and Execution of the Budget, Circular No. A-11 (2020).

NOTES

[18] Tax provisions were categorized by level of complexity based on recordkeeping intensity, tax planning activities, and overall difficulty of extracting information from the taxpayer's financial books (IRS 2023b).

[19] "Instructions for the 1040," Internal Revenue Service, accessed on July 17, 2024, https://www.irs.gov/instructions/i1040gi.

[20] Yellen, Janet, "Memorandum for Commissioner Rettig: IRS Operational Plan," September 15 2022, https://www.taxnotes.com/tax-notes-today-federal/tax-system-administration/yellen-requests-irs-plan-resource-implementation/2022/08/18/7dym5?highlight=yellen%20rettig.

[21] An IRS spokesperson told the reporter that they could not explain the discrepancy.

[22] The denominator includes abandoned calls, disconnects, and busy signals.

[23] Data from the Treasury Inspector General for Tax Administration (TIGTA) indicate that only 63 percent of additional taxes recommended by examiners in operational audits in fiscal years 2015 through 2019 were ultimately assessed (after administrative appeals and abatements; TIGTA 2021). That figure is likely even lower due to further reductions on judicial review.

[24] In some cases, the statute of limitations is extended beyond three years. Those exceptions include (1) failure to file a required tax return; (2) agreement between the IRS and taxpayer to extend the period; (3) taxpayer reported less than 25 percent of their income on the tax return; and (4) the taxpayer filed a false or fraudulent tax return with intent to avoid taxes. The three-year time limit can be suspended if the IRS issued a notice of deficiency (with the IRS's proposed assessment) or the taxpayer filed for bankruptcy.

[25] Philip Swagel, "The effect of increased funding for the IRS," *CBO Blog*, Congressional Budget Office, September 2, 2021, https://www.cbo.gov/publication/57444.

[26] Secretary of the Treasury Janet Yellen, Letter to Ways and Means Chairman Richard Neal, September 14, 2021. https://home.treasury.gov/system/files/136/Yellen_Neal_Congressional_Budget.pdf.

[27] Natasha Sarin, "The Case for a Robust Attack on the Tax Gap," Featured Stories, US Department of the Treasury,. September 7, 2021, https://home.treasury.gov/news/featured-stories/the-case-for-a-robust-attack-on-the-tax-gap.

REFERENCES

American Customer Service Index. 2023. *American Customer Service Index Federal Government Report 2023*. Ann Arbor, MI: American Customer Service Index.

Anderson, Brandon, Keegan Brown, Damon Frezza, Zane Girouard, Alissa Graff, Tom Hertz, Navya Kambalapally, Piper Kurtz, Kara Leibel, Justin Nave, Mark Payne, Daniel Rodriguez, and Brian Sartain. 2024. "Research on Race and Ethnicity: 2024 Update." Presentation given at the 14th Annual IRS/TPC Joint Research Conference on Tax Administration, Washington, DC, June 13.

Auten, Gerald, and David Splinter. 2024. "Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends." *Journal of Political Economy* 132 (7): 2179–227. https://www.journals.uchicago.edu/doi/full/10.1086/728741.

Berger, Daniel, Eric Toder, Victoria Bryant, John Guyton, and Patrick Langetieg. 2018. *Estimating the Effects of Tax Reform on Compliance Burden*. Washington, DC: Urban-Brookings Tax Policy Center.

Boning, William, Nathaniel Hendren, Ben Sprung-Keyser, and Ellen Stuart. 2023. "A Welfare Analysis of Tax Audits across the Income Distribution." Working Paper 31376. Cambridge, MA: National Bureau of Economic Research.

CBO (Congressional Budget Office) 2014. *How Initiatives to Reduce Fraud in Federal Health Programs Affect the Budget*. Washington, DC: CBO.

------. 2020. *Trends in the Internal Revenue Service and Enforcement Funding*. Washington, DC: CBO.

DeBacker, Jason, Bradley Heim, Ahn Tran, and Alexander Yuskavage. 2020. "Tax Noncompliance and Measures of Income Inequality." *Tax Notes Federal*: 1103-1118. https://www.taxnotes.com/tax-notes-federal/compliance/tax-noncompliance-and-measures-income-inequality/2020/02/17/2c3y5

Derby, Elena and Connor Dowd and Jacob Mortenson. 2024. "Statistical Bias in Racial and Ethnic Disparity Estimates Using BIFSG." Washington, DC: Joint Committee on Taxation.

Elzayn, Hadi, and Evelyn Smith, Thomas Hertz, Arun Ramesh, Robin Fisher, Daniel Ho, and Jacob Goldin. 2023. "Measuring and Mitigating Racial Disparities in Tax Audits." Working Paper. Palo Alto, CA: Stanford Institute for Economic Policy Research.

GAO (General Accounting Office). 2000. "IRS Is Working to Improve Its Estimates of Compliance Burden." GAO/GGD-01-11. Washington, DC: GAO.

------. 2024. "Tax Gap: IRS Should Take Steps to Ensure Continued Improvements in Estimates." GAO-24-106449. Washington, DC: GAO.

Guyton, John, and Ronald Hodge II. 2013. "The Compliance Costs of IRS Post-Filing Processes." Papers Given at the 2013 IRS-Tax Policy Center Research Conference, Urban Institute, Washington, DC, June 20.

Guyton, John and Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman. 2023. "Tax Evasion at the Top of the Income Distribution: Theory and Evidence." Washington, DC: Internal Revenue Service.

Guyton, John and Kara Leibel, Dayanand Manoli, Ankur Patel, Mark Payne, and Brenda Schafer. 2024. "The Effects of EITC Correspondence Audits on Low-Income Earners." In *Tax Policy and the Economy*, vol. 38, edited by Robert Moffitt, 163–207. Chicago, Illinois: University of Chicago Press.

Hemel, Daniel, Janet Holtzblatt, and Steven Rosenthal. 2022. *The Tax Gap's Many Shades of Gray*. Washington, DC: Urban-Brookings Tax Policy Center.

Hertz, Tom, Brian Sartain, Kara Leibel, and Mark Payne. 2023. "Differences in Audit Rates by Race." Presentation given at the 13th Annual IRS/TPC Joint Research Conference on Tax Administration, Washington, DC, June 22.

HM Revenue and Customs. 2020. "Non-detection Multipliers for Measuring Tax Gaps." Working Paper. London, UK: HM Revenue and Customs.

Hodge, Ronald II, Alan Plumley, Kyle Richison, Getaneh Yismaw, Nicole Nisek, Matt Olson, and H, Sanith Wijesinghe. 2015. "Estimating Marginal Revenue/Cost Curves for Correspondence Audits." Presentation given at the 2015 IRS/TPC Research Conference on Tax Administration, Washington, DC, June 18.

REFERENCES

Holtzblatt, Janet, and Jamie McGuire. 2016. "Factors Affecting Revenue Estimates of Tax Compliance Estimates." Working Paper 2016-5. Washington, DC: Congressional Budget Office.

------. 2020. *Effects of Recent Reductions in the Internal Revenue Service's Appropriations on Returns on Investment*. Washington DC: Urban-Brookings Tax Policy Center.

IRS (Internal Revenue Service). 2022a. *Comprehensive Taxpayer Attitude Survey 2021 Executive Report*. Publication 5296 (Rev. 4-2022). Washington D.C.: IRS.

------. 2022a. *Internal Revenue Service Data Book, 2021*. Publication 55-B. Washington, DC: IRS.

------. 2022b. *Tax Gap Estimates for Tax Years 2014-2016*. Publication 1415 (Rev. 10-2022). Washington, D.C.: IRS.

------. 2023a. *Taxpayer Compliance Burden*. Publication 5743 (Rev. 4-2023). Washington, DC: IRS.

------. 2023b. *IRS Inflation Reduction Act Strategic Operating Plan*. Publication 3744 (Rev. 4-2023). Washington, D.C.: IRS.

------. 2024a. *Fiscal Year 2025 Congressional Budget Justification and Annual Performance Report and Plan*. Publication 4450 (Rev. 2-2024). Washington, DC: IRS.

------. 2024b. *Return of Investment: Re-Examining Revenue Estimates for IRS Funding*. Publication 5901 (2-2024). Washington, DC: IRS.

------. 2024c. *Internal Revenue Service Data Book, 2023*. Publication 55-B. Washington, DC: IRS.

------. 2024d. *IRA Strategic Operating Plan: Annual Update Supplement*. Publication 3744-A (4-2024). Washington, DC: IRS.

Johns, Andrew, and Joel Slemrod. 2010. "The Distribution of Income Taxes Noncompliance." *National Tax Journal* 63 (3): 397-418. https://www.journals.uchicago.edu/doi/abs/10.17310/ntj.2010.3.01.

Marcuss, Rosemary, George Contos, John Guyton, Patrick Langetieg, Allen Lerman, Susan Nelson, Brenda Schafer, and Melissa Vigil. 2013. "Income Taxes and Compliance Costs: How Are They Related?" *National Tax Journal* 68 (4): 833–53. https://www.journals.uchicago.edu/doi/10.17310/ntj.2013.4.03.

National Taxpayer Advocate. 2023. *Annual Report to Congress*. Washington, DC: Taxpayer Advocate Service.

Rifaat, Alexander. 2024. "Conflicting IRS Level-of-Service Figures Raise Questions." *Tax Notes*. 183. April 22, https://www.taxnotes.com/tax-notes-today-federal/tax-system-administration/conflicting-irs-level-service-figures-raise-questions/2024/04/16/7jf9n.

Sarin, Natasha, and Mark Mazur. 2024. "The Inflation Act's Impact on Tax Compliance and Fiscal Sustainability." *Tax Notes*. February 19. https://www.taxnotes.com/featured-analysis/inflation-reduction-acts-impact-tax-compliance-and-fiscal-sustainability/2024/02/16/7j509.

Sarin, Natasha, and Lawrence Summers. 2019. "Shrinking the Tax Gap: Approaches and Revenue Potential." *Tax Notes*. November 18. https://www.taxnotes.com/special-reports/compliance/shrinking-tax-gap-approaches-and-revenue-potential/2019/11/15/2b47g

TIGTA (Treasury Inspector General for Tax Administration). 2021. *Trends in Compliance Activities through Fiscal Year 2019*. Report Number 2021-30-011. Washington, DC: TIGTA.

------. 2023. *Final Results of the 2023 Filing Season*. Report Number 2024-400-006. Washington, DC: TIGTA.

------. 2024. *Quarterly Snapshot: The IRS's Inflation Reduction Act Spending Through March 31, 2024*. Report Number 2024-IE-R015. Washington, DC: TIGTA. https://www.tigta.gov/sites/default/files/reports/2024-06/2024ier015fr.pdf

US Department of the Treasury. 2021. *The American Families Plan Tax Compliance Agenda*. Washington, DC: Treasury.

Werfel, Daniel. 2023. "The Filing Season and the IRS Budget." Written Testimony before the US House Ways and Means Committee, Washington, DC, April 27.

ABOUT THE AUTHOR

Janet Holtzblatt is a senior fellow at the Urban-Brookings Tax Policy Center. Before joining the Tax Policy Center, Holtzblatt was the unit chief for tax policy studies in the Tax Analysis Division of the Congressional Budget Office. Before that, she was the deputy director of the Individual Taxation Division in the US Treasury Department's Office of Tax Analysis and a senior analyst on the Senate Budget Committee. Holtzblatt's research focuses on racial disparities in the income tax, the earned income tax credit, marriage penalties and bonuses in the income tax system, the administration of the tax code, and tax simplification. Holtzblatt served on the board of directors for the National Tax Association. In 2017, she received the Referee of the Year Award from the National Tax Journal. She was the recipient of the 2020 Davie-Davis Public Service Award, which is bestowed annually by the National Tax Association. Holtzblatt earned a BA in economics and history from the University of Illinois and a PhD in economics from the University of Wisconsin.

The Tax Policy Center is a joint venture of the
Urban Institute and Brookings Institution.

BROOKINGS

For more information, visit taxpolicycenter.org
or email info@taxpolicycenter.org